

Collegio Carlo Alberto



Gender Wage Gaps Reconsidered: A Structural
Approach Using Matched Employer-Employee
Data

Cristian Bartolucci

No. 116

September 2009 (Revised, July 2010)

Carlo Alberto Notebooks

www.carloalberto.org/working_papers

GENDER WAGE GAPS RECONSIDERED: A STRUCTURAL APPROACH USING MATCHED EMPLOYER-EMPLOYEE DATA^{*}

Cristian Bartolucci
Collegio Carlo Alberto[†]

July 2010

Abstract

In this paper I propose and estimate an equilibrium search model using matched employer-employee data to study the extent to which wage differentials between men and women can be explained by differences in productivity, disparities in friction patterns, segregation or wage discrimination. The availability of matched employer-employee data is essential to empirically disentangle differences in workers productivity across groups from differences in wage policies toward those groups. The model features rent splitting, on-the-job search and two-sided heterogeneity in productivity. It is estimated using German microdata. I find that female workers are less productive and more mobile than males. Female workers have on average slightly lower bargaining power than their male counterparts. The total gender wage gap is 42 percent. It turns out that most of the gap, 65 percent, is accounted for by differences in productivity, 17 percent of this gap is driven by segregation while differences in destruction rates explain 9 percent of the total wage-gap. Netting out differences in offer-arrival rates would increase the gap by 13 percent. Due to differences in wage setting, female workers receive wages 9 percent lower than male ones.

JEL Code: J70, C51, J64

KEYWORDS: Labor market discrimination, search frictions, structural estimation, matched employer-employee data.

^{*}This article is a revised version of a chapter from my PhD dissertation at CEMFI. I am specially grateful to Manuel Arellano for his guidance and constant encouragement. I would also like to thank Stéphane Bonhomme, Christian Bontemps, Raquel Carrasco, Zvi Eckstein, Chris Flinn, Andrés Erosa, Pietro Garibaldi, Carlos González-Aguado, José Labeaga, Joan Llull, Claudio Michelacci, Pedro Mira, Enrique Moral-Benito, Diego Puga, Jean Marc Robin, Nieves Valdez and seminar participants at Uppsala University, Bank of Canada, Collegio Carlo Alberto, Bank of Italy, Bank of France, PSE, UQAM, CEMFI, University of Tucuman, University College London, European Meeting of the Econometric Society (2009), Cosme Workshop, EC-Squared Conference on Structural Microeconometrics, SEA meetings (2008) and IZA-ESSLE (2008) for very helpful comments. Special thanks are due to Nils Drews, Peter Jacobebbinghaus and Dana Muller at the Institute for Employment Research for invaluable support with the data.

[†]Via Real Collegio, 30 - 10024 - Moncalieri (Torino), Italy. Email: cristian.bartolucci@carloalberto.org.

1 Introduction

In this paper I propose and estimate an equilibrium search model using matched employer-employee data to study the extent to which wage differentials between men and women can be explained by differences in productivity, disparities in friction patterns, segregation or wage discrimination. The model features rent-splitting, on-the-job search and two-sided heterogeneity in productivity. The estimation involves several steps: firstly, I estimate group-specific productivity from firm-level production functions. Secondly, I compute job-retention and job-finding rates using employee-level data. Finally, I calculate rent-splitting parameters (bargaining power) relying on individual wage data, transition parameters and productivity measures estimated in the previous stages.

There has been many studies focused on explaining how much of the unconditional mean wage differential between groups may be understood as wage discrimination¹. The traditional approach takes the unexplained gap in wage regressions as evidence of discrimination. This method involves estimating Mincer-type equations for both groups and then decomposing the difference in mean wages into “explained” and “unexplained” components. The fraction of the gap that cannot be explained by differences in observable characteristics is considered to be discrimination. This kind of analysis has been very informative from a descriptive perspective, but the causal interpretation and the nature of discrimination are not clear.

Discrimination refers to differences in wages that are only caused by the fact of belonging to a given group. Therefore causality is an essential issue in this context. Ideally, detecting discrimination would require testing whether the group effect is significant once we have controlled for between-groups differences in wage determinants.

The availability of matched employer-employee data allowed Hellerstein and Neumark (1999)² to pioneer a new approach. Their method uses firm

¹See Blau and Kahn (2003) and Altonji and Blank (1999) for good surveys.

²The main papers in this branch are: Hellerstein and Neumark, (1999) with Israeli data, Hellerstein, Neumark and Troske, (1999) with U.S. data, Verner (1999) using data from Ghana, Crepon, Deniau and Pérez-Duarte (2002) with French data, Lopez-Acevedo

level data to estimate relative marginal products of various worker types, which are then compared with their relative wages. This analysis implies a clear causality from productivity to wages. Whenever perfect competition holds in the labor market, productivity is the only wage determinant, and therefore any difference in wages that is not driven by a difference in productivity may be considered to be discrimination.

However, a frictionless scenario has been shown to be a little help in understanding the labor market. In a labor market with frictions, productivity is not the only wage determinant, and therefore comparing wages and productivity may provide an incomplete picture of the problem. Moreover, wage differentials across groups are often accompanied by unemployment rate and job duration differentials. There is a vast literature estimating differentials in job-finding and job-retention rates across groups, directly observing duration in unemployment and employment or using experiments in audit studies. Although there is agreement in predicting effects of frictions on wages³, there is scarce empirical evidence on how much of the wage gap can be accounted for by differences in friction patterns.

Estimated structural models may provide a complete interpretation of observed wage gaps as a consequence of between-groups differences in wage determinants. Nevertheless, progress in this direction has been slow mainly due to the difficulty in separately identifying the impacts of skill differentials and discrimination from worker-level survey data. The main references are Eckstein and Wolpin (1999) and Bowlus and Eckstein (2002). Both papers are focused on racial discrimination in the U.S. and deal with this empirical identification problem through structural assumptions. Eckstein and Wolpin (1999) proposed a method based on a two-sided, search-matching model that formally accounts for unobserved heterogeneity and unobserved offered wages. They argued that differences in the bargaining power parameter (their index of discrimination) are not identified unless some firm-side data are available, and hence they were forced to simply compute bounds

et al (2005) with data from Mexico, Kawaguchi (2007) with Japanese data, Zhang and Dong (2009) with Chinese data, Van Biesebroeck, (2009) with data from three Subsaharian countries.

³See van den Berg and van Vuuren (2003) for a good discussion on this issue.

for discrimination that ended up being not informative on the estimation sample they worked with. Bowlus and Eckstein (2002) also proposed a search model with heterogeneity in workers productivity but including an appearance-based employer disutility factor. As long as there are firms that do not discriminate, their method is able to identify between-groups differences in the skill distribution as well as the discrimination parameter, which in their case, was the proportion of discriminatory employers⁴. The focus of Bowlus and Eckstein (2002) was not on estimation, as the main objective was to propose an identification strategy⁵.

The first attempt to use an equilibrium search model to study gender discrimination was made by Bowlus (1997). In her paper, Bowlus only focused on the effect of gender differences in friction patterns on wage differentials without distinguishing between differences in productivity and discrimination. In a recent paper Flabbi (2010) used the identification strategy proposed in Bowlus and Eckstein (2002), but allowing for heterogeneity in matches productivity⁶. The model was estimated by maximum likelihood to study whether gender labor market differentials are due to labor market discrimination or to unobserved productivity differences.

In this paper I propose and estimate an equilibrium search model with on-the-job search, rent-splitting, and productivity heterogeneity in firms and workers. The availability of matched employer-employee data furthers identification by allowing me to disentangle differences in workers productivity across groups from differences in wage policies toward those groups. I combine productivity measures estimated at the firm level *a la* Hellerstein and Neumark (1999), group-specific friction patterns estimated from individual duration data, and individual wages to estimate the wage equation provided by the structural model. This structural wage equation states the precise

⁴Mondal (2006) also estimates a similar model to study racial wage differentials in the U.S.

⁵The model assumes no firm heterogeneity and generates counterfactual implications on wage distributions. They are only able to match some moments generated by the model with moments estimated using a sample from the NLSY.

⁶In order to have a model that is estimable with employee-level data, Flabbi (2010) only includes heterogeneity at the match level and does not allow for on-the-job search. Although he allows for wage bargaining, the bargaining power is not estimated.

relationship between wages, workers ability, firms productivity, friction patterns, and bargaining power.

Most of the variables included in standard wage equations, such as education or experience may be understood as proxies of the true wage determinants, that are the worker productivity, the outside options, and the rent splitting rule. The wage equation presented in this paper may be understood as the structural counterpart of a standard Mincer-equation, but only including *theoretically relevant* wage determinants⁷. It allows me to undertake counterfactual analysis, such as comparing wages of two ex-ante identical workers in terms of productivity and outside options, who only differ in the rent-splitting parameter corresponding to their gender⁸.

The approach presented in this paper links two independent branches of the wage-discrimination literature. It is a natural step in extending the structural estimation literature focused on wage discrimination. Due to the availability of better data, I am able to estimate a more complete model⁹, but also, it allows me to use a more robust identification strategy to measure differences in productivity and to have clean measures of the effect on wages of labor market segregation. However, this approach may also be understood as an evolution of the approach proposed by Hellerstein and Neumark (1999) where only productivity gaps were estimated. Here I am providing a complete interpretation of the observed wage gap, where the productivity gap is only one of the potential determinants of the difference in wages between males

⁷The structural wage equation may be also understood as an equation that completes the Hellerstein et al (1999) approach, where wages were assumed to be simply equal to productivity. The model provide a close form solution in which wages are found to be a function of productivity but also function of friction patterns and bargaining power.

⁸Note that a difference in the bargaining power between men and women is considered as wage discrimination. This has already been assumed in Eckstein and Wolpin (1999) and it is meaningful in the sense that an inequality in the rent-splitting parameter generates a difference in wages between two workers with the same ability and outside option that are working in similar jobs in terms of sector and qualification and they only differ in terms their gender.

⁹Particularly related with Flabbi (2010), that at the moment is the most complete structural estimation in the wage-discrimination literature, the model presented in this paper allows for two-sided heterogeneity and on-the-job search. In the estimation, rent splitting parameters are not imposed but estimated and I am not forcing the same firms distribution across gender. See sections 2 and 4 for details.

and females.

I use a 1996-2005 panel of matched employer-employee data provided by the German Labor Agency, called LIAB¹⁰. This dataset is especially useful for the current study for two reasons. Firstly, it contains essential individual variables such as gender, wages and occupation. Secondly, it is a panel that tracks firms as opposed to individuals, which is important for estimating production functions using panel estimation methods. To the best of my knowledge, this paper presents the first structural estimation using matched employer-employee data to study labor market discrimination¹¹.

The empirical analysis proceeds by first calculating differences in productivity between men and women, following the approach in Hellerstein and Neumark (1999). As in this study, I find important negative productivity differentials for women. Next, I analyze group-specific dynamics. I find that women have higher job-creation rates than equivalent men, and that females also have higher job-destruction rates than males. Finally, I estimate group-specific bargaining power. In spite of having large wage differentials, on average, women are only found have slightly lower bargaining power than men.

In terms of wages, the total gender wage gap is 41 percent. It turns out that most of the gap, 65 percent, is accounted for by differences in productivity. Differences in destruction rates explain 9 percent while differences in the distribution of firm's productivity faced by male and female workers explain 17 percent of the total wage-gap. Netting out differences in offer-arrival rates would increase the gap by 13 percent. Differences in the rent-splitting parameter are responsible for 21 percent of the wage gap, which implies that female workers receive wages 8.6 percent lower than equivalent males.

The rest of the paper is organized as follows. In the next section, I describe the structural model. Section 3 describes the data. In Section 4, I present the

¹⁰This dataset is subject to strict confidentiality restrictions. It is not directly available but only after the IAB has approved the research project, The Research Data Center (FDZ) provides on site use or remote access to external researchers.

¹¹There is a recent paper by Sulis (2007) that studies gender wage differentials in Italy, estimating a structural model. Sulis uses employee level data with firm identifiers, without data on firms, such as capital or output.

estimation of the structural model inputs, namely the productivity measures and friction parameters, I present and discuss these intermediate results, and finally I describe the empirical strategy to estimate the structural wage equation and its results. In Section 5, the counterfactual experiments are discussed and I compare my empirical results with those resulting from other strategies for detecting discrimination using the same data. Conclusions are offered in Section 6.

2 Structural framework

In this section I describe the behavioral model of labor market search with matching and rent-splitting. The main goal of estimating a structural model is to clearly state a wage setting equation that allows me to measure the effect of each wage determinant. Having this wage equation estimated, it is straightforward to obtain the effect of discriminatory wage policies, comparing a man's wage with the wage that a woman with the same wage determinants would receive.

Previous research has shown the ability of this kind of models in describing the labor market outputs and dynamics. Building on these assessments, in this paper I am interested in using the structural model as a measurement tool that allows me to empirically disentangle the effect of each wage determinant on the gender wage gap. Search-matching models has been used as an instrument to address empirical questions in a variety of papers. Examples are the previously mentioned papers in the discrimination literature, but there are also interesting contributions in measuring returns to education (Eckstein and Wolpin, 1995) or in analyzing the effect of a change in the minimum wage (Flinn, 2006).

This paper presents a quantitative exercise that aims to measure the empirical relevance of wage discrimination. As in Eckstein and Wolpin (1999), the model does not take an explicit stance on what the sources of wage discrimination are. Rather, I intend to measure if firm's payments differ if the employee is male or female, *ceteris paribus*.

An alternative approach is presented in Flabbi (2010). He is also in-

interested in measuring discrimination, but he assumes that discrimination comes from employer's tastes. As in Eckstein and Bowlus (2002), employer heterogeneity in tastes for discrimination, has a fundamental role in his identification strategy. Without firm level data, the existence of employers that do not discriminate is essential to identify differences in productivity. Having matched employer employee data allows me to avoid such kind of assumptions.

Taking a stand about what type of discrimination is generating the data, is important because it opens up the possibility for policy experiments. But policy experiments are more interesting whenever wage discrimination is empirically more relevant. Therefore the first question that has to be answered is whether firms set different wages by gender.

2.1 Assumptions

I propose a continuous time, infinite horizon, stationary economy. This economy is populated by infinitely lived firms and workers. All agents are risk neutral and discount future income at rate $\rho > 0$.

Workers: I normalize the measure of workers to one. Workers may belong to one of different groups (k) defined in terms of gender¹². Workers have different abilities (ε) measured in terms of efficiency units they provide per unit of time. The distribution of ability in the population of workers is exogenous and specific to each group, with cumulative distribution function $L_k(\varepsilon)$. Group-specific distributions of efficiency units provided by workers is crucial to consider between groups differences in productivity. This source of heterogeneity is perfectly observable by every agent in the economy. Each worker may be either unemployed or employed. The workers from a generic group k that are not actually working receive a flow utility, proportional to their ability, $b_k\varepsilon$.

Firms: Every firm is characterized by its productivity (p). I assume that p is distributed across firms according to a given cumulative distribution

¹²The structural model abstracts many dimensions that may be relevant in the wage setting. In order to compare jobs as similar as possible, the empirical analysis is clustered at sector and occupation level. See section 4 for details.

function $H_k(p)$, which is continuously differentiable with support $[p_{\min}, p_{\max}]$. The distribution of firms's productivity is group-specific, what allows the model to be robust to labor market segregation of workers groups. This source of heterogeneity is perfectly observable by every agent in the economy. The opportunity cost of recruiting a worker is zero.

Each firm contacts a worker of a given group k at the same constant rate, regardless of the firm's bargained wage, its productivity or how many filled job it has. Unemployed workers receive job offers at a Poisson rate $\lambda_{0k} > 0$. Employed workers may also search for a better job while employed and they receive job offers at a Poisson rate $\lambda_{1k} > 0$. I treat λ_{0k} and λ_{1k} as exogenous parameters specific to each group k . Searching while unemployed as searching while employed has no cost. Employment relationships are exogenously destroyed at a constant rate $\delta_k > 0$, leaving the worker unemployed and the firm with nothing. The marginal product of a match between a worker with ability ε and a firm with productivity p is εp .

Whenever an employed worker meets a new firm, the worker must choose an employer and then, if she switches employers, she bargains with the new employer with no possibility of recalling her old job. If she stays at her old job, nothing happens. Consequently when a worker negotiates with a firm, her alternative option is always the unemployment. The surplus generated by the match is split in proportions β_k and $(1 - \beta_k)$, for the worker and the firm respectively, where $\beta_k \in (0, 1)$, is exogenously given and specific to each group k . I will refer to β_k as the rent-splitting parameter. As in Wolpin and Eckstein (1999), I interpret $\beta_{male} - \beta_{female}$ as an index of the level of discrimination in the labor market. A difference in β in the same kind of job and sector, reveals differential payments unrelated to productivity and outside options, which are only driven by belonging to a given group.

It is well known that many other factors may have an impact on the rent-splitting parameters. the main candidates are, negotiation skills, risk aversion and the discount factor. In this model agents are assumed to be risk neutral and the discount rate is homogenous across groups. Although this could be part of the story that explains differences in β , there are scarce convincing empirical evidence of gender differences in these primitive para-

meters¹³. This could be the reason why risk aversion and the discount factor have been held constant across genders in most of the empirical studies on wage discrimination.

The model assumes that the worker does not have the option of recalling the old employer, there is no possibility of Bertrand competition between firms as in Cahuc, Postel-Vinay and Robin (2006). Whether to allow firm competition a la Bertrand or not is controversial. While the Cahuc et al bargaining scenario may be conceptually more appealing and may help to avoid the Shimer critique, it is not clear how realistic this assumption is. Mortensen (2003) argues that counteroffers are uncommon empirically, and Moscarini (2008) shows that, in a model with search effort, firms may credibly commit to ignore outside offers to their employees, letting them go without a counteroffer, and suffer the loss, in order to keep in line the other employees' incentives to not search on the job. Moreover, it can be shown that including a marginally positive cost of negotiation, it will not be profitable for firms to try to poach the worker to better firms, and then the Bertrand competition vanishes.

In an environment where contracts cannot be written and wages are continuously negotiated, the alternative option of the job is always unemployment. In this context, if a worker receives an offer from a firm with higher productivity, she must switch. She cannot use this offer to renegotiate with her current firm, because she knows that tomorrow this offer will not be available and then her future option will be the unemployment again¹⁴. This possibility is also discussed in Flinn and Mabli (2010).

Beyond the theoretical relevance of between-firms Bertrand competition, this assumption is not critical for most of the results presented in this paper.

¹³The main contributions come from experiments, see Croson and Gneezy (2009) for a good survey.

¹⁴If wages are continuously negotiated, firms could increase the wage of the worker at the moment of the on-the-job offer to try to avoid the worker quitting. If the alternative employer is more productive it can force the transition by also paying a premium. This auction for the worker finishes when the actual firm cannot pay more than the full productivity and transition holds as in a Bertrand competition. This premium may be considered as a hiring cost for the firm. Modelling this possibility is outside the scope of this paper.

In the appendix, I show that using the same data with a variation of the model where between-firm Bertrand competition is allowed, the gender wage-gap decomposition remains practically unchanged.

It is not clear whether β can be interpreted as a Nash bargaining power. Shimer (2006) argues that in a simple search-matching model with on-the-job search, the standard axiomatic Nash bargaining solution is inapplicable, because the set of feasible payoffs is not convex. This non-convexity arises because an increase in the wage has a direct negative effect over the firm's rents but an indirect positive effect raising the duration of the job. This critique will hold out depending on the shape of the productivity distribution. Whether β can be understood as a Nash Bargaining Power, is not essential for this study. If this critique holds up, I still interpret β as a rent-splitting parameter that simply states the proportion of the surplus that goes to the worker. A difference in these parameters remains informative about wage-discrimination.

This model is similar to the model presented in independent work by Flinn and Mabli (2010). The main difference is in the distributions of productivity. In order to have a model that is estimable with employee-level data only, they assume that there is a technologically-determined discrete distribution of worker-firm productivity. In other words, they assume discrete heterogeneity at the match level while here I assume two-side continuous heterogeneity. The model presented here also have the convenient property of producing a closed form solution for the wage setting equation.

2.2 Value Functions

The expected value of income for a worker with ability ε , who belongs to group k , currently employed at wage $w(p, \varepsilon, k)$ is denoted by $E(w(p, \varepsilon, k), \varepsilon, k)$ and it satisfies:

$$\begin{aligned}
& \rho E(w(p, \varepsilon, k), \varepsilon, k) \\
= & w(p, \varepsilon, k) + \delta_k (U(\varepsilon, k) - E(w(p, \varepsilon, k), \varepsilon, k)) + \\
& \lambda_{1k} \int_{w(p, \varepsilon, k)}^{w(p, \varepsilon, k)_{\max}} [E(\tilde{w}(p, \varepsilon, k), \varepsilon, k) - E(w(p, \varepsilon, k), \varepsilon, k)] dF(\tilde{w}(p, \varepsilon, k)|\varepsilon, k).
\end{aligned} \tag{1}$$

The expected value of being unemployed for a worker with ability ε , who belongs to group k is given by:

$$\begin{aligned}
& \rho U(\varepsilon, k) \\
= & b_k \varepsilon + \lambda_{0k} \int_{w(p, \varepsilon, k)_{\min}}^{w(p, \varepsilon, k)_{\max}} [E(\tilde{w}(p, \varepsilon, k), \varepsilon, k) - U(\varepsilon, k)] dF(\tilde{w}(p, \varepsilon, k)|\varepsilon, k).
\end{aligned}$$

Finally, the value of the match with productivity $p\varepsilon$ for the firm when paying a wage $w(p, \varepsilon, k)$ to a worker of group k is given by:

$$\begin{aligned}
& \rho J(w(p, \varepsilon, k), p, \varepsilon, k) \\
= & p\varepsilon - w(p, \varepsilon, k) - (\delta_k + \lambda_{1k} \bar{F}(w(p, \varepsilon, k)|\varepsilon, k)) J(w(p, \varepsilon, k), p, \varepsilon, k),
\end{aligned} \tag{2}$$

where $\bar{F}(w(p, \varepsilon, k)|\varepsilon, k) = 1 - F(w(p, \varepsilon, k)|\varepsilon, k)$ and $F(w(p, \varepsilon, k)|\varepsilon, k)$ is the equilibrium cumulative distribution function of wages paid by firms with productivity lower than p to workers with ability ε who belongs to group k . Note that every parameter is group-specific. As the alternative value of the match for the firm is always zero, this value does not depend on alternative matches and therefore it is independent of parameters of the other groups of workers. Although every group is sharing the same labor market, all the value functions may be considered group by group as if they were in independent markets. For notation simplicity I then omit the k -index. Note that $w(p, \varepsilon)$ is a function of p and ε , therefore given p and ε , the wage is a redundant state variable which is only included for exposition simplicity.

These expressions are equivalent to the value functions of the model with heterogenous firms presented in Shimer (2006) including heterogeneity in

workers ability. But here, wages are determined by the following surplus splitting rule:

$$(1 - \beta) [E(w(p, \varepsilon), \varepsilon) - U(\varepsilon)] = \beta J(w(p, \varepsilon), p, \varepsilon). \quad (3)$$

After some algebra (see the appendix for the proof), it can be shown that:

$$\begin{aligned} w(p, \varepsilon) &= p\varepsilon - (\rho + \delta + \lambda_1 \bar{F}(w(p, \varepsilon)|\varepsilon)) \\ &\quad \times \frac{(1 - \beta)}{\beta} \int_{w(p, \varepsilon)_{\min}}^{w(p, \varepsilon)} \frac{1}{(\rho + \delta + \lambda \bar{F}(\tilde{w}(p, \varepsilon)|\varepsilon))} d(\tilde{w}(p, \varepsilon)). \end{aligned}$$

Noting that $\bar{F}(w(p, \varepsilon)|\varepsilon) = \bar{H}(p)$ and changing the variable within the integral, I obtain a first-order differential equation,

$$w(p, \varepsilon) = p\varepsilon - (\rho + \delta + \lambda_1 \bar{H}(p)) \frac{(1 - \beta)}{\beta} \int_{p_{\min}}^p \frac{1}{(\rho + \delta + \lambda \bar{H}(p'))} \frac{d(w(p, \varepsilon))}{dp'} dp'.$$

Solving the differential equation, after some algebra the wage equation takes the following form:

$$w(p, \varepsilon) = \varepsilon p - \varepsilon(1 - \beta)(\rho + \delta + \lambda_1 \bar{H}(p))^\beta \int_{p_{\min}}^p (\rho + \delta + \lambda_1 \bar{H}(p'))^{-\beta} dp'. \quad (4)$$

This expression states a clear relationship between wages $w(p, \varepsilon)$, workers' ability (ε), firm productivity (p), friction patterns (λ_1, δ) and the rent-splitting parameter (β). This wage equation is relatively similar to the one proposed by Cahuc, Postel-Vinay and Robin (2006) when the wage is bargained between a firm with productivity p and an unemployed worker with ability ε ¹⁵.

As expected the model predicts that the mean equilibrium wage increases in β , and that the mean wage paid by a firm with productivity p increases in p . Note in (4) that, if $\beta = 1 \Rightarrow w(p, \varepsilon) = p\varepsilon$, the maximum wage that a firm

¹⁵Note that in Cahuc et al (2006) when the wage is bargained between a firm and an unemployed worker the Bertrand competition does not hold and therefore their proposed scenario is equivalent to this one. The only difference comes from the fact that both parts take into account future Bertrand competition.

with productivity p can pay to a worker with ability ε is the full productivity. If $\beta = 0 \Rightarrow w(p, \varepsilon) = p_{\min}\varepsilon$, that is the minimum wage that a worker would accept to leave unemployment, see Figure 1¹⁶.

As it can be seen in Figure 1, the mean equilibrium wage increases when λ_1 increases and when δ decreases. Many models in the literature predict that the mean equilibrium wage decreases in the amount of frictions (see for example the models in Burdett and Mortensen, 1998, Bontemps, Robin and van den Berg, 2000, Postel-Vinay and Robin, 2002 and Cahuc, Postel-Vinay and Robin 2006). The intuition behind this fact is clearly explained in van den Berg and van Vuuren (2003). They argue that all of these models are asymmetric in workers and employers. This asymmetry is due to the fact that workers correspond to a relatively long-lived unit whereas firms can expand and contract, and can be created and destroyed relatively quickly. When frictions decrease, the value of creating a vacancy increases, and this may prompt an instantaneous inflow of new firms. The latter mitigates the effect of the reduction in frictions on the firms whereas it increases the effect on the workers, and hence the wage increases.

I have assumed that the economy is in steady state. The standard stationary equilibrium conditions are exploited. The inflow must balance the outflow for every stock of workers, defined in terms of individual ability, employment status and, for those workers who are employed, firm's productivity.

- The inflow to the unemployment must be equal to its outflow, $\lambda_0\mu = \delta(1 - \mu)$, where μ is the unemployment rate given by:

$$\mu = \frac{\delta}{\delta + \lambda_0} \quad (5)$$

- The inflow to jobs in firms with productivity p or lower than p must be equal to its outflow:

$$\lambda_0 H(p)\mu = (\lambda_1 \bar{H}(p) + \delta) G(p)(1 - \mu),$$

¹⁶These simulations are calibrated using the estimated parameters of male skilled workers in the manufacturing sector, see Section 4. Those parameter are: $\beta = 0.292$, $\lambda_1 = 0.217$ and $\delta = 0.034$.

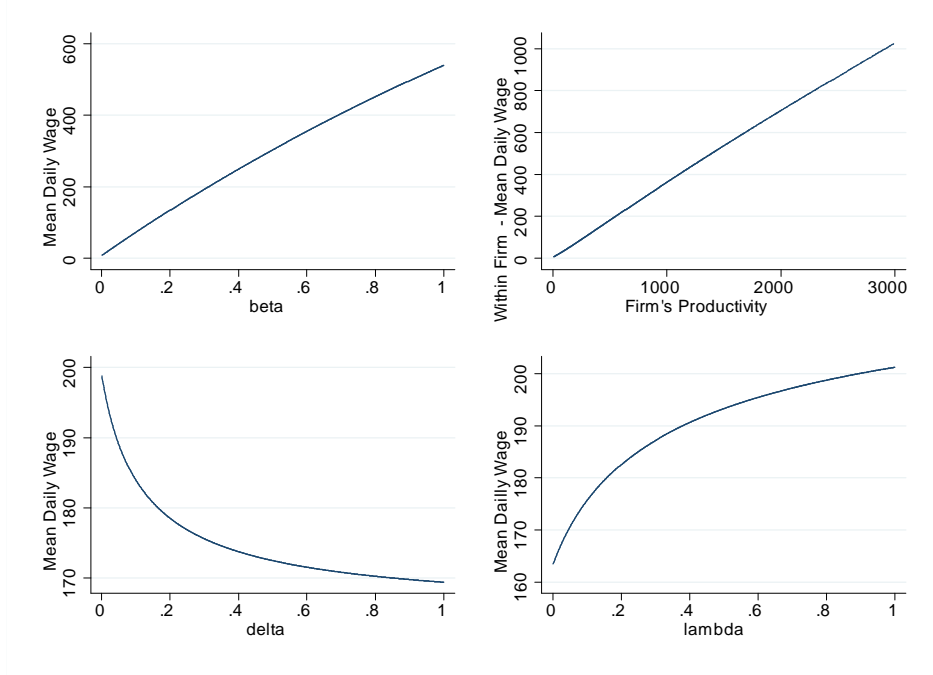


Figure 1: Wage Setting Equation

where $G(p)$ is the fraction of workers employed at a firm with productivity p or lower than p . Then using condition (5) and rearranging:

$$G(p) = \frac{H(p)}{1 + \kappa_1 \bar{H}(p)}, \quad (6)$$

where κ_1 is $\frac{\lambda_1}{\delta}$. This stationary condition, (or its counterpart in terms of wages) is quite common and has been broadly used after Burdett and Mortensen (1998) to infer the primitive distribution of productivity (or the primitive distribution of wages) when only the distribution of productivity (or distribution of wages) within employed workers is observable. Since here I use matched employer-employee data, I can directly observe the empirical distribution of productivity at firm level. I only use this stationary condition in order to construct the likelihood for the duration analysis in section 4.

- The fraction of employed workers with ability ε or lower than ε that are working in firms with productivity p or lower than p are $(1 - \mu)\tilde{F}(\varepsilon, p)$,

where $\tilde{F}(\varepsilon, p)$ is the joint cdf of ε and p . These workers leave this group due to a better offer or because they become unemployed, such event occurs with probability $(\delta + \lambda_1 \bar{H}(p))$. The inflow to this group is given by the unemployed workers with ability ε or lower than ε , (*ie*: $L(\varepsilon)\mu$) who receive an offer from a firm with productivity p or lower than p . This last event occurs with probability $\lambda_0 H(p)$. Then I have the following condition:

$$(1 - \mu)(\delta + \lambda_1 \bar{H}(p))F(\varepsilon, p) = \lambda_0 H(p)L(\varepsilon)\mu.$$

Next using conditions (5) and (6), and rearranging:

$$F(\varepsilon, p) = \frac{H(p)}{(1 + \kappa_1 \bar{H}(p))} L(\varepsilon) = G(p)L(\varepsilon). \quad (7)$$

This expression says that *there is no sorting between firm's productivity and worker's ability*¹⁷.

This statement is controversial, and there is an active debate in the assortative matching literature about it. Becker (1973) showed that in a model without search frictions but with transferable utility, if there are supermodular production functions, any competitive equilibrium exhibits positive assortative matching. In more recent work, Shimer and Smith (2000) and Atakan (2006) show that in search models, complementarities in production functions are not sufficient to ensure assortative matching. Assuming different cost functions the first one predicts a negative correlation while the second one the opposite.

After Abowd, Kramarz, and Margolis (1999), the empirical literature has mainly focused on estimating the correlation between worker's and firm's fixed effects using matched employer-employee data. However, there are still no definitive results. Abowd et al found a negative and small correlation between firms and workers fixed effects for France, and zero correlation for the U.S. while Lindeboom, Mendez and van den Berg (2010), using a Portuguese

¹⁷To show that there is no sorting, condition (7) is necessary but not sufficient. We also need that the p_{\min} , the minimum productivity, is independent of the worker ability. This condition also holds in this model (the proof is in the appendix).

matched employer-employee dataset, find that there is positive assortative matching.

3 Data

Linked Employer-Employee Data from the German Federal Employment Agency or LIAB.

I use the linked employer-employee dataset of the IAB (denoted LIAB) covering the period 1996-2005. LIAB was created by matching the data of the IAB establishment panel and the process-produced data of the Federal Employment Services (Social security records). The distinctive feature of this data is the combination of information about individuals with details concerning the firms in which these people work. The workers source contains valuable data on age, sex, nationality, daily wage (censored at the upper earnings limit for social security contributions¹⁸), schooling/training, the establishment number and occupation based on a 3-digit code that in this paper is collapsed into two categories: skilled and unskilled jobs.¹⁹

The firm's data give details on total sales, value added, investment, depreciation²⁰, number of workers and sector²¹. In particular, only firms with more than 10 workers, positive output and positive depreciated capital have been included in my subsample. Since firms of different sectors do not share the same market, I construct separate samples for each sector. LIAB has a very detailed industry classification. I focus on four main industries: Manufacturing, Construction, Trade, and Services²². Participation of establishments is

¹⁸In the sample of firms used in this paper, 14.7% of the worker observations have censored information on wages. This proportion varies significantly across gender and occupation. 4.7% of female observations and 18.1% of males observations are censored. The proportion of worker observations in high-qualification occupations, with wages that exceed the upper earning limit for social security contributions is 37.0% while the corresponding one to low-qualification occupations is 3.1%.

¹⁹I have assigned the following groups to the unskilled category: Agrarian occupations, manual occupations, services and simple commercial or administrative occupations. While I have classified as skilled jobs: Engineers, professional or semi-professional occupations, qualified commercial or administrative occupations, and managerial occupations.

²⁰The survey gives information about investment made to replace depreciated capital.

²¹For a more detailed description of this dataset, see Alda et al (2005)

²²The service sector includes three kind of services defined in the survey: industrial

voluntary, but the response rates are high, exceeding 70 per cent. However, the response rate in some key-variables for my purpose is lower. Among survey respondent, only 60% of firms in the previous four industries provide valid responses for output. To estimate productivity I need data on output and number of workers in each category. I only consider observations from the old Federal Republic of Germany (West-Germany). Finally, firms with strictly less than 10 employees were removed. The final number of observations in my sample of firms is 15,174. Table 1 provides descriptive statistics of the final sample of firms.

Table 1: Firms - Descriptive Statistics

	n° of firms	Output (mean)*	n° of workers	Women (%)		Men (%)	
				Unsk.	Skill	Unsk.	Skill
MANUFACT.	7,354	151.0	4,297,762	11.9	8.1	55.3	24.7
CONSTRUCT.	1,491	30.2	170,786	12.9	11.5	58.1	17.6
TRADE	2,078	67.4	247,884	30.6	17.5	34.4	17.6
SERVICES	4,251	30.4	1,043,678	21.1	21.0	38.9	20.0
TOTAL	15,174	92.8	5,760,110	14.2	11.0	51.5	23.4

* Per annum total output in millions of euros

One of the main advantages of this data-set is that it has information on all the employees subject to social security in each firm²³. The employee data are matched to firms for which I have valid estimates of productivity through a unique firm identifier. The raw data contains 21,246,022 observations between 1996 and 2005, but after this final trimming I have a 9-year unbalanced panel, including a total of 5,760,110 workers' observations distributed into 15,174 firms' observations.

services, transport and communication, and other services.

²³Employees subject to social security are workers, other employees and trainees who are liable to health, pension and/or unemployment insurance or whose contributions to pension insurance is partly paid by the employer. The following forms of employment are not considered liable to social security: civil servants, self-employed persons, unpaid family workers and so-called "marginal" part-time workers (A "marginal" part-time worker is a person who is either: employed only short-term or paid a maximum wage of €400 per month).

Given this selection, the sample becomes less representative. According to the Federal Statistical Office (Statistisches Bundesamt Deutschland), between 1996 and 2005 the proportion of the workforce in the manufacturing sector ranges between 25.6 and 31.7 percent, while in the sample used in this paper it exceeds 74%, see Table 1. For that reason, when making inference on the total population, between-groups aggregation of results is made using the group weights obtained from GSOEP²⁴.

In the sample women are, on average, younger than men, they have less tenure and less experience. Women tend to have high-skill occupations with higher frequency than men. The proportion of immigrants is higher within the men's group. See Table 2 for details of the workers sample.

Table 2: Workers - Descriptive Statistics

	WOMEN	MEN
IMMIGRANT (%)	8.4	10.4
AGE (YEARS)	39.2	40.7
TENURE (YEARS)	10.1	12.0
EXPERIENCE (YEARS)	15.3	17.1
SKILLED (%)	46.4	31.9
OBSERVATIONS	1,290,156	4,130,453

The main goal of this study is to understand the gender wage gap. The difference in conditional means is 21 percent (see Table 17 in the appendix). Meaning that women, on average, have salaries 21 percent lower than men with the same observable characteristics. The unconditional wage differential averages 42 percent, but it is not stable across sectors and occupations (see

²⁴The weights of each groups are estimated with the relative frequencies in the 1996-2005 sample of the GSOEP, which are: manufacturing-skilled-men 12.3%, manufacturing-unskilled-men 15.6%, manufacturing-skilled-women 6.1%, manufacturing-unskilled-women 4.8%, construction-skilled-men 4.1%, construction-unskilled-men 7.0%, construction-skilled-women 1.1%, construction-unskilled-women 0.2%, trade-skilled-men 6.1%, trade-unskilled-men 2.9%, trade-skilled-women 10.5%, trade-unskilled-women 2.8%, services-skilled-men 10.3%, services-unskilled-men 4.2%, services-skilled-women 8.7%, services-unskilled-women 3.1%.

Table 3). Mean-wages estimated across industries and occupations show that the gap ranges between 30 percent and 45 percent. Wage gaps are significantly different from zero in every sector and in every group, and they are larger for skilled workers in manufacturing, construction and trade sectors.

Table 3: Gender Wage Gap

		MEAN DAILY-WAGE		W-GAP (%)
		WOMEN	MEN	
MANUFACT.	UNSKILLED	76.94 (0.41)	109.25 (0.59)	29.57% (1.00%)
	SKILLED	105.74 (0.11)	190.30 (0.22)	44.43% (0.33%)
CONSTRUCT.	UNSKILLED	53.35 (0.14)	96.59 (0.16)	44.77% (0.30%)
	SKILLED	82.81 (0.27)	150.90 (0.54)	45.11% (0.81%)
TRADE	UNSKILLED	47.64 (0.08)	84.90 (0.53)	43.16% (0.51%)
	SKILLED	67.76 (0.02)	119.20 (0.09)	43.88% (0.11%)
SERVICES	UNSKILLED	49.11 (0.07)	88.47 (0.17)	44.80% (0.24%)
	SKILLED	87.44 (0.26)	156.31 (0.68)	44.06% (0.94%)
WEIGHTED AVERAGE		72.08	124.10	41.93%

Note: Standard errors are given in parentheses. Means of log-wage are estimated using worker-level data maximizing saturated normal-likelihoods at the firm level. Means of wages are calculated by the moment generating function. Standard errors are obtained by Delta-Method.

German Socio-Economic Panel

The LIAB version used in this paper is a panel of firms complemented with workers data. As it does not track workers, it is not possible to dis-

tinguish between attrition²⁵ and job-termination²⁶. For that reason I use GSOEP (German Socio-Economic Panel) to estimate group-specific transition parameters²⁷. The German Socio-economic panel is a representative repeated survey of households in Germany. This survey has been carried out annually with the same people and families in Germany since 1984 (but I only use 1996-2005)²⁸.

4 Empirical Strategy and Results

The discrete nature of annual data implies a complicated censoring of the continuous-time trajectories generated by the theoretical model. Because of these complications a potentially efficient full information maximum likelihood is not considered as a candidate for the estimation. Instead, I perform a multi-step estimation procedure²⁹.

Even though it may be theoretically inefficient, I prefer a step-by-step method. One reason is that the efficiency of full information maximum likelihood is only guaranteed in the case of correct specification. However I am interested in having productivity differences and transition parameter estimates that are robust to misspecification in other parts of the model. Another reason is that transition parameters are better estimated using a standard labor force survey such as GSOEP.

A multi-step estimation procedure allows me to have control of the source of variation that is effectively identifying each parameter. The empirical identification of productivity differences with firm level data is weak and imprecise. Full-information maximum likelihood may have helped empirically

²⁵There is no attrition in a establishment, which is the unit of observation in the sample. I lack individuals that may have changed their identifier or that have changed establishment without changing firm.

²⁶Unless the worker leaves the establishment and moves to another establishment within the panel.

²⁷Cahuc, Postel-Vinay and Robin (2006) follow the same strategy for estimating transition parameters with the French Labor Force Survey.

²⁸See Wagner, Burkhauser, and Behringer (1993) for further details on the GSOEP.

²⁹Multi-step estimation has been done in many papers. Good examples are Bontemps, Robin, and Van den Berg (2000), Postel-Vinay and Robin (2002), and Cahuc, Postel-Vinay and Robin (2006).

because data on wages may be used to improve on the productivity estimates, but on the other hand I would not be able to guarantee that such estimates are solely revealing productivity differences as opposed to wage setting inequalities. If the model were the true data generating process this caveat would not be necessary, because the model does not imply any reverse causality from wages to productivity, and the noise in productivity estimates would be only due to the contemporary productivity shock uncorrelated with wages. However, even in an informal way, models are generally incomplete and it seems prudent to use estimators that are as robust to misspecification as possible.

The structural model abstracts many dimensions that may be relevant in the wage setting, for example amenities or union pressure. These omitted dimensions may be mainly associated with different types of jobs. As it can be seen in Tables 1 and 2 there are important differences between men and women in terms of occupation and sector. In order to compare jobs which are as similar as possible, the empirical analysis is clustered at the sector and occupation level. The model is estimated independently for each of the four sectors. In order to control for occupation; transition parameters and the rent-splitting parameter are also estimated independently for both types of jobs, in each sector and gender group. I only control for occupation parametrically when I estimate productivity, because I need to consider the full workforce in each firm.

4.1 Productivity

The production function specification chosen in the empirical section, is a standard Cobb-Douglas function with constant returns to scale and quality adjusted labor input. This function has already been used in the discrimination literature to estimate between-group productivity differences and it is also consistent with the theory proposed in the previous section. The value added, Y_{jt} , produced by firm j in period t , is given by:

$$Y_{jt} = A_j K_{jt}^{(1-\alpha)} L_{jt}^{\alpha} e^{u_{jt}}$$

Where K_{jt} is the total capital, A_j is a firm-specific productivity parameter, u_{jt} is a zero mean stationary productivity shock and L_{jt} is the total amount of labor in efficiency units given by:

$$L_{jt} = \sum_K \tilde{\gamma}_k L_{jt}^k$$

As it was previously mentioned, I have four types of workers depending on gender (men and women) and occupation (skilled and unskilled). I normalize $\gamma_{ms} = 1$ considering male skilled workers as the reference group³⁰. Now $\gamma_k = \tilde{\gamma}_k / \tilde{\gamma}_{ms}$ is the proportional productivity of group k relative to the productivity of male skilled workers. Imposing constant returns to scale and assuming that firms can adjust capital instantaneously makes this specification totally consistent with the theory, where I have assumed that the productivity of a match is $p\varepsilon$. Section A.3 in the appendix provides more details and robustness checks on this assumption.

Using the panel with firm level data on value-added³¹ (Y_{jt}), depreciated capital³² (K_{jt}^d) and number of workers in each category, I estimate the production function in logs forcing constant returns to scale and constant proportionality between occupation across gender ($\gamma_{wu} = \gamma_w \times \gamma_u$).

$$\begin{aligned} \log(Y_{jt}) = & \log(A_j) + (1 - \alpha_l) \log(K_{jt}^d) + \\ & \alpha_l \log(L_{jt}^{ms} + \gamma_w L_{jt}^{ws} + \gamma_u L_{jt}^{mu} + \gamma_w \gamma_u L_{jt}^{wu}) + u_{jt} \end{aligned} \quad (8)$$

where L_{jt}^{ws} and L_{jt}^{ms} are, respectively, the number of women and men in skilled occupations in firm j at time t while L_{jt}^{wu} and L_{jt}^{mu} are, respectively, the number of women and men in unskilled occupations in firm j at time t .

³⁰Due to this normalization, the firm specific productivity \tilde{A}_j is redefined as $A_j \gamma_{ms}^{\alpha_l}$.

³¹I only use value-added in manufacturing. Output measures are used in construction, trade and services due to lack of convergence of estimates based on value-added. Assuming that a constant fraction of output is spent in materials, both types of estimates are consistent for the same parameters, the difference goes to the constant term. In order to use firm productivity measures in the structural wage equation both measures are not equivalent because the constant term matters, and hence, value-added is used in every sector.

³²Assuming that a constant fraction (d) of capital depreciates by unit of time: $K_{jt}^d = d \times K_{jt} \Rightarrow \log(K_{jt}^d) = \log(d) + \log(K_{jt})$. Therefore $\alpha_k \log(d)$ goes to the constant term.

The model predicts that more productive firms are able to attract more workers of every type. As a result the total labor input would be correlated with the firm fixed effect. Therefore, I estimate (8) by Within-Groups Non Linear Least Squares to remove the firm fixed effects.

Table 4: Production Function Estimates

WG-NLLS ESTIMATES OF (8)			
	α_l	γ_w	γ_u
MANUFACTURING	0.963 (0.005)	0.672 (0.062)	0.484 (0.042)
CONSTRUCTION	0.961 (0.006)	0.701 (0.052)	0.444 (0.039)
TRADE	0.971 (0.009)	0.804 (0.092)	0.487 (0.056)
SERVICES	0.945 (0.007)	0.588 (0.068)	0.298 (0.030)
WEIGHTED AVERAGE	0.96	0.67	0.43

Note: Time dummies included. Robust Standard errors are given in parentheses. Weighted averages take into account the number of firms in each sector.

The Non-Linear Within Groups results are shown in Table 4. Women's productivity is lower than men's productivity in similar jobs. This difference ranges between 20 percent and 41 percent. On average across cells, female workers are 33% less productive than male ones in each job. One of the main candidates to explain this large productivity gap is that these estimates are not taking into account that women work on average less hours than men. Using the GSOEP³³, I find that the average hour-gap is 19.9 percent, hence differences in hours are likely to be one of the main determinants of the productivity-gap³⁴.

Unskilled workers are also found to be between 51 percent and 70 percent less productive than skilled workers. As in most production function

³³LIAB does not provide information on hours.

³⁴Although differences in hours are shown to be important, the main results of this paper remain valid. I only mention them in order to have a better understanding of the estimated productivity gap. See Section 5.1 for a more detailed discussion on this issue.

estimations using microdata, α_l is found to be very near one and hence, α_k is very small but statistically different from zero. Although this finding is standard³⁵, the main results of this paper are not very sensitive to this issue. If instead of $\alpha_l = 0.96$, I include an *a priori* more realistic value of $\alpha_l = 0.60$, the wage gap decomposition does not change significantly.

Pioneered by Hellerstein and Neumark (1999) and Hellerstein, Neumark and Troske (1999), differences in productivity across gender are now well documented in the literature. The first paper finds, with Israeli firm-level data, a productivity gap of 17 percent while the second, using a U.S. sample of manufacturing plants reports a productivity gap of 15 percent. These studies have been criticized mainly due to the potential endogeneity of the proportion of female workers in the firm³⁶. In this paper, I treat the number of workers of each group as potentially correlated with the firm fixed effect³⁷. Estimating (8) by Within-Groups Non Linear Least-Squares the firm fixed effect is completely removed, hence my estimates are robust to any correlation of the labor input level and the labor input composition with the firm fixed effect.

In this dataset there is strong evidence of correlation between the firm's fixed effect and the firm's labor input. Estimating (8) by NLLS without fixed effects, γ 's estimates are significantly lower, the average of γ_w^{NLLS} across sectors is 0.38 and the average of γ_u^{NLLS} across sectors is 0.26. See Table 12 in the appendix.

Estimating (8) by non-linear within-groups the firm fixed effect is removed, but the simultaneity problem is not totally solved. One alternative would be to treat L_{jt} and K_{jt} as predetermined variables and to estimate the production function by Non Linear GMM. This possibility has been attempted, but there is a severe problem of lack of precision on the GMM estimates of the γ parameters.

³⁵One interpretation of this result is that K_{jt}^d only captures variable capital whereas fixed capital is subsumed in the firm effect. But if so, the constant returns restriction is dubious

³⁶See Altonji & Blank (1999).

³⁷Indeed, the model predicts that more productive firms are able to attract more workers of every type. Therefore, the total labor input will be correlated with the firm fixed effect but not the labor input composition.

The precision in the non-linear GMM estimates of γ 's is low in every sector, also using with different sets of instruments. γ has been estimated using: Only lagged levels for the equation in differences as in Arellano and Bond (1991); lagged levels for the equation in differences and lagged differences to instrument the equation in levels as in Arellano and Bover (1995) and only lagged differences to instrument the equation in levels as in Cahuc et al. (2006). These three alternative sets of moment conditions have been used treating the proportion of each kind of worker as endogenous, but also as exogenous and the estimated γ remained imprecise.

As productivity differentials end up being the main cause of the wage gaps, I was particularly concerned about the robustness of their estimation. I obtained extremum estimators that minimizes the two-stage robust GMM2 objective function and iterated-GMM but also Chernozhukov and Hong (2003) MCMC type of estimators for Continuously-updated GMM. Non-Linear System GMM and NLLS estimates of the production function are reported in section A.3 in the appendix³⁸.

The lack precision in the quality parameter estimates is a pervasive problem in this kind of production function specification. In Cahuc, Postel-Vinay and Robin (2006), the authors have finally decided to estimate the productivity parameters and the wage equation parameter simultaneously by an iterated non-linear least squares procedure without removing the firm's fixed effect.

4.2 Labor Market Dynamics

Given that job terminations occur due to a job-to-job transitions or to the exogenous job destruction, and that both processes are Poisson, the model defines the precise distribution of job durations t conditional on the firm productivity p :

$$\mathcal{L}(t|p) = [\delta + \lambda_1 \bar{H}(p)] e^{-[\delta + \lambda_1 \bar{H}(p)]t}. \quad (9)$$

³⁸MATA codes for computing the non-linear estimators previously described are available from the author upon request.

As I use GSOEP to estimate transition parameters and this dataset does not have productivity measures, λ_1 and δ are estimated treating p as an unobservable. Therefore, I maximize the unconditional likelihood $\mathcal{L}(t) = \int \mathcal{L}(t|p)g(p)dp$, where $g(p)$ is the probability density function of firm's productivity among employed workers.

Taking derivatives with respect to p in equation (6), I get the density of firm's productivity in the population of workers:

$$g(p) = \frac{(1 + \kappa_1)h(p)}{1 + \kappa_1\bar{H}(p)} \quad (10)$$

In the appendix I show the individual contribution to the unconditional likelihood becomes simple enough to be estimated and it is given by:

$$\mathcal{L}(t) = \frac{\delta(1 + \kappa_1)}{\kappa_1} \left[\int_{\delta t}^{(1+\kappa_1)\delta t} \frac{e^{-x}}{x} dx \right]$$

Integrating unobserved productivity out of the conditional likelihood removes p and all reference to the sampling distribution $H(p)$ (Cahuc et al, 2006). This method is robust to any misspecification in the wage bargaining. The only property of the structural model that is required, is that there exist a scalar firm index, in this case p , which monotonously defines transitions.

In the appendix, I show how to obtain the exact form of the likelihood that takes into account that some durations are right-censored while some others started before the survey's beginning. Finally, an individual contribution to the log-likelihood is:

$$\begin{aligned} l(t_i) = & (1 - c_i) \log \left(\frac{\int_{\delta t}^{(1+\kappa_1)\delta t} \frac{e^{-x}}{x} dx}{\frac{e^{-\delta H_i}}{\delta} - \frac{e^{-\delta(1+\kappa_1)H_i}}{\delta(1+\kappa_1)} - H_i \int_{\delta H_i}^{(1+\kappa_1)\delta H_i} \frac{e^{-x}}{x} dx} \right) + \\ & c_i \log \left(\frac{\frac{e^{-\delta t_i}}{\delta} - \frac{e^{-\delta(1+\kappa_1)t_i}}{\delta(1+\kappa_1)} - t_i \int_{\delta t_i}^{(1+\kappa_1)\delta t_i} \frac{e^{-x}}{x} dx}{\frac{e^{-\delta H_i}}{\delta} - \frac{e^{-\delta(1+\kappa_1)H_i}}{\delta(1+\kappa_1)} - H_i \int_{\delta H_i}^{(1+\kappa_1)\delta H_i} \frac{e^{-x}}{x} dx} \right) \end{aligned}$$

Where c_i is a right-censored spell indicator and H_i is the time period elapsed before the sample started³⁹.

³⁹The MATA code for computing the exponential integral and the MATA code to max-

Table 5: Transition Parameters - Maximum Likelihood Estimates

	UNSKILLED					
	WOMEN			MEN		
	λ_1	δ	κ_1	λ_1	δ	κ_1
MANUFACT.	0.406	0.044	9.202	0.314	0.031	10.095
	(0.041)	(0.004)	(0.127)	(0.021)	(0.002)	(0.176)
CONSTRUCT.	0.601	0.098	6.085	0.437	0.105	4.162
	(0.188)	(0.031)	(0.150)	(0.025)	(0.006)	(0.118)
TRADE	0.5257	0.094	5.613	0.432	0.074	5.478
	(0.050)	(0.009)	(0.085)	(0.042)	(0.008)	(0.090)
SERVICES	0.559	0.095	5.866	0.458	0.086	5.313
	(0.051)	(0.009)	(0.073)	(0.037)	(0.007)	(0.049)
	SKILLED					
	WOMEN			MEN		
	λ_1	δ	κ_1	λ_1	δ	κ_1
MANUFACT.	0.308	0.044	7.025	0.217	0.034	6.373
	(0.031)	(0.004)	(0.316)	(0.015)	(0.002)	(0.103)
CONSTRUCT.	0.510	0.090	5.691	0.255	0.071	3.620
	(0.095)	(0.017)	(0.107)	(0.025)	(0.006)	(0.048)
TRADE	0.327	0.060	5.459	0.353	0.050	7.093
	(0.020)	(0.004)	(0.078)	(0.032)	(0.004)	(0.219)
SERVICES	0.393	0.073	5.363	0.227	0.050	4.547
	(0.024)	(0.004)	(0.062)	(0.015)	(0.003)	(0.051)

Note: Per annum estimates. Standard errors are given in parentheses.

Maximum likelihood estimates are reported in Table 5. The average duration of an employment spell, $1/\delta$ (possibly changing employer) is between 10 and 32 years, but the mean-duration across sectors is 20.2 years. The average time between two outside offers, $1/\lambda_1$, ranges from 1.7 to 4.6 years. These results seem to be fairly large but they are compatible with others in the literature. van den Berg and Ridder (2003) using a similar specification but with German aggregated data, find δ equal to 0.060 and κ_1 equal to 6.5⁴⁰, while here a weighted average of δ across sectors and groups is 0.0574

imize this likelihood are available from the author upon request.

⁴⁰van den Berg and Ridder (2003, p.237) report monthly rates for $\lambda_1 = 0.028$ and $\kappa_1 = \frac{\lambda_1}{\delta} = 6.5$.

and the weighted average of κ_1 is 6.41.

Skilled workers have in general lower transition rates to unemployment and lower on-the-job offer arrival rates. Women are more mobile than men in terms of job-to-job transitions and, in general, they also have higher job-destruction rates. Considering κ_i as an index of frictions it is noteworthy that in general, women and unskilled workers suffer higher labor market frictions than men and skilled workers respectively.

4.3 The Wage Equation: Closing the Model

The structural wage equation (4) can be written as:

$$w_{j,t,i} = \varepsilon_i w_{j,t,k(i)}^p(p_{j,t}, \beta_{k(i)}, H_{k(i)}(p), \lambda_{1k(i)}, \delta_{k(i)})$$

where $w_{j,t,i}$ is the daily wage of a worker i , who belongs to a group $k(i)$, in a firm j with productivity p_j at time t , and:

$$\begin{aligned} & w_{j,t,k(i)}^p(p_{j,t}, \beta_{k(i)}, H_{k(i)}(p), \lambda_{1k(i)}, \delta_{k(i)}) \\ = & p_{j,t} - (1 - \beta_{k(i)})(\rho + \delta_{k(i)} + \lambda_{1,k(i)} \bar{H}_{k(i)}(p_{j,t}))^{\beta_{k(i)}} \\ & * \int_{p_{\min}}^{p_{j,t}} (\rho + \delta_{k(i)} + \lambda_{1,k(i)} \bar{H}_{k(i)}(p'))^{-\beta_{k(i)}} dp' \end{aligned}$$

As shown in equation (7) ε is statistically independent of p , thus

$$\begin{aligned} E(w_{j,t,i}) &= E(\varepsilon_i w_{j,t,k(i)}^p(p_{j,t}, \beta_{k(i)}, H_{k(i)}(p), \lambda_{1k(i)}, \delta_{k(i)})) \\ &= E_k(\varepsilon) E(w_{jtk}^p(p_{j,t}, \beta_{k(i)}, H_{k(i)}(p), \lambda_{1k(i)}, \delta_{k(i)})) \end{aligned}$$

$E_k(\varepsilon) = \gamma_k$ is the mean efficiency units of workers in group k in that market relative to the male skilled group. Therefore the predicted mean wage for workers of group k working in firms with productivity $p_{j,t}$ at time t is:

$$E(w_{jtk}) = \gamma_{k(i)} w_{jtk}^p(p_{j,t}, \beta_{k(i)}, H_{k(i)}(p), \lambda_{1k(i)}, \delta_{k(i)}) \quad (11)$$

The group chosen for normalization is unimportant. Changing this group to a generic group k , we would change our measure of productivity. Instead

of p_j , that is, the productivity measured in terms of efficiency units of skilled males, we would have $p_j^k = \gamma_k p_j$, that is, the productivity measured in term of efficiency units of group k . In fact, to define (11) in terms of the productivity of group k , we only need to put γ_k inside the expectation operator:

$$\begin{aligned} E(w_{j,t,i}) &= E(\gamma_{k(i)} p_{j,t} - \\ &\quad (1 - \beta_{k(i)})(\rho + \delta_{k(i)} + \lambda_{1,k(i)} \bar{H}_{k(i)}(p_{j,t}))^{\beta_{k(i)}} \\ &\quad * \int_{p_{\min}}^{p_{j,t}} (\rho + \delta_{k(i)} + \lambda_{1,k(i)} \bar{H}_{k(i)}(p'))^{-\beta_{k(i)}} \gamma_k dp') \end{aligned}$$

Noting that $\frac{dp}{dp^k} = \frac{1}{\gamma_k}$, and changing the variable within the integral, we have

$$\begin{aligned} E(w_{j,t,i}) &= E(p_{j,t}^k - \\ &\quad (1 - \beta_{k(i)})(\rho + \delta_{k(i)} + \lambda_{1,k(i)} \bar{H}_{k(i)}(p_{j,t}^k))^{\beta_{k(i)}} \\ &\quad * \int_{p_{\min}^k}^{p_{j,t}^k} (\rho + \delta_{k(i)} + \lambda_{1,k(i)} \bar{H}_{k(i)}(p^{k'}))^{-\beta_{k(i)}} dp^{k'}) \quad \blacksquare \end{aligned}$$

For each firm in the sample I estimate the average daily wage \bar{w}_{jtk} paid to workers of group k at time t . Since wages are top-coded, I estimate the firm mean-wage for each worker group (*ie* : \bar{w}_{jtk}) by maximum likelihood at the firm level assuming that wages are log-normal⁴¹. Under the steady state assumption and according to the theory presented in section 2, \bar{w}_{jtk} exhibits stationary fluctuation around the steady state mean wage $E(w_{jtk})$ paid by firm j with productivity p_j .

I estimate equation (11) in logs with firm-level data, that is to estimate

⁴¹Wages are linear in ε and there is not other source of within firm variation in wages. Therefore, the within firm distribution of wages is the same than the distribution of ability. I am assuming log-normality in the distribution of ε .

$$\begin{aligned}
\log \bar{w}_{jtk} = & \ln(\gamma_k) + \\
& \ln \left(p_{j,t} - (1 - \beta_{k(i)})(\rho + \delta_{k(i)} + \lambda_{1,k(i)} \bar{H}_{k(i)}(p_{j,t}))^{\beta_{k(i)}} \right. \\
& * \left. \int_{p_{\min}}^{p_{j,t}} (\rho + \delta_{k(i)} + \lambda_{1,k(i)} \bar{H}_{k(i)}(p'))^{-\beta_{k(i)}} dp' \right) \\
& + v_{jtk},
\end{aligned} \tag{12}$$

by weighted non-linear least squares at the firm level, where γ_k , δ_k and λ_k are parameters estimated in previous stages, $p_{j,t}$ is the productivity⁴² of firm j at time t and v_{jtk} is a transitory shock with unrestricted variance. As usual, the discount factor has been set to an annual rate of 5% (daily rate of 0.0134%).

Standard errors have to take into account that γ , λ_1 and δ are estimated in previous stages. To solve this problem I combine bootstrap for γ with the analytical solution for λ_1 and δ . Hence, I obtain standard errors replicating the productivity estimation and the bargaining power estimation in 200 resamples of the LIAB original sample, with replacement, but taking the transition parameters as the population ones. To correct these preliminary standard errors, I add to them the analytical term corresponding to the standard errors of λ_1 and δ reported in Table 5. Finding the analytical solution is not difficult in this case because estimators come from different samples so that I can omit the term corresponding to the outer product of scores in the first and second stages.

Consistent standard errors are given by:

$$\begin{aligned}
\widehat{Var}(\hat{\beta}) = & Var(\hat{\beta}|\hat{\lambda}_1, \hat{\delta})_{bootstrap} + \\
& \frac{\hat{H}_{\beta\lambda_1} \widehat{Var}(\hat{\lambda}_1) \hat{H}_{\beta\lambda_1} + \hat{H}_{\beta\delta} \widehat{Var}(\hat{\delta}) \hat{H}_{\beta\delta}}{\hat{H}_{\beta\beta} * \hat{H}_{\beta\beta}}
\end{aligned}$$

where H is the objective function in the optimization, which in this case is the weighted sum of squares and $H_{\xi\phi} = \frac{\partial(\frac{\partial H}{\partial \xi})}{\partial \phi}$. Second derivatives of H are

⁴²See section A.4 in the appendix for details about how I recover p_{jt} using parameters estimated in previous stages. In section A.4 I also present robustness checks on different assumptions regarding the construction of p_{jt} .

obtained numerically⁴³.

Results are presented in Table 6. Women are found to have lower rent-splitting parameters than men in construction and trade for both skilled and unskilled occupations, and in manufacturing skilled occupations. Female workers receive larger portion of the surplus than males in services and in manufacturing unskilled occupations but these differences are not significant.

Table 6: Rent-Splitting Parameter Estimates

WNLLS ESTIMATES OF (12)			
		WOMEN	MEN
		β	β
MANUFACTURING	UNSKILLED	0.419 (0.129)	0.398 (0.096)
	SKILLED	0.226 (0.088)	0.292 (0.036)
CONSTRUCTION	UNSKILLED	0.214 (0.090)	0.408 (0.045)
	SKILLED	0.113 (0.078)	0.186 (0.104)
TRADE	UNSKILLED	0.339 (0.173)	0.382 (0.106)
	SKILLED	0.152 (0.066)	0.222 (0.064)
SERVICES	UNSKILLED	0.849 (0.125)	0.757 (0.125)
	SKILLED	0.413 (0.104)	0.324 (0.073)

Note: Corrected Bootstrap Standard errors are given in parentheses.

There is a clear pattern in terms of skilled and unskilled occupations. Unskilled workers receive larger shares of the surplus in every sector, considering female and male workers⁴⁴. These findings are not consistent with the

⁴³The analytical correction $\frac{\hat{H}_{\beta\lambda_1} \widehat{Var}(\hat{\lambda}_1) \hat{H}_{\beta\lambda_1} + \hat{H}_{\beta\delta} \widehat{Var}(\hat{\delta}) \hat{H}_{\beta\delta}}{\hat{H}_{\beta\beta} * \hat{H}_{\beta\beta}}$ is not significant in any industry either for skilled or like the unskilled workers.

⁴⁴These differences, as differences between industry may be understood as consequences

results found in Cahuc, Postel-Vinay and Robin (2006), where they report a positive association between bargaining power and job qualification.

Estimates of the rent-splitting parameter are considerably higher than the ones reported on Cahuc, Postel-Vinay and Robin (2006). This is probably due to differences in our definition of match rents⁴⁵. In a similar model estimated with US employee-level data by Flinn and Mabli (2010), the overall bargaining power is found to be 0.45 while here the weighted average across cells is remarkably similar, 0.421.

Allowing between-firms Bertrand competition as in Cahuc, Postel-Vinay and Robin (2006), changes the magnitude of β , but not the gender differences. In the appendix, I show a numerical exercise, where β are recovered by the simulated method of moments using the same data and a model with between-firms Bertrand competition. The bargaining power is found to be significantly lower, the weighted average is 0.219 in this case, but the gender and occupation patterns do not change. Women are found to have, lower bargaining power than males in the construction and trade, while there is not a clear pattern in manufacturing and services. As in the model without Bertrand competition, workers in low qualification occupations are also found to have higher β than workers in high qualification occupations, which suggests that this occupation pattern is not a modelling artifact, but only a difference between German and French labor markets⁴⁶.

Differences in rent-splitting parameters are not significant in every sector. I only find that male workers receive larger shares of the surplus than female ones in the construction sector, where bootstrap p-values of the differences in β are 95.5% for unskilled workers and 90.4% for skilled ones.

of compensating differentials or differences in union pressure. They cannot be understood as discrimination because we are comparing occupations and not workers.

⁴⁵The surplus is defined in terms of the productivity of the match and the outside option. Both models imply different outside options. Without Bertrand competition the worker outside option is the unemployment. While allowing for Bertrand competition, the worker outside option is the whole productivity of the poaching firm. As the outside option in the model with Bertrand competition dominates the one in the model without Bertrand competition, the estimated bargaining power is smaller.

⁴⁶See Section A.5 for details.

5 Wage-Gap Decomposition

The structural wage setting equation provides us with a direct way of isolating the effect of each wage determinant over the overall wage differential. Hence, we are able to calculate which fraction of the wage gap is due to segregation or differences in the rent-splitting parameter, productivity or friction patterns.

Using the structural wage equation (4):

$$w_{j,t,i} = \varepsilon_i w^p(p_{j,t}, \beta_{k(i)}, H_{k(i)}(p), \lambda_{1k(i)}, \delta_{k(i)})$$

for each sector and each worker group it is possible to measure the wage differential caused by differences in each wage determinant. For example I can estimate the wage gap accounted by β as the relative difference between the mean-wage that women actually receive and the mean wage that women would receive if they had the male rent-splitting parameter.

$$1 - \frac{E[\varepsilon_i w^p(p_{j,t}, \beta_W, H_W(p), \lambda_{1,W}, \delta_W)]}{E[\varepsilon_i w^p(p_{j,t}, \beta_M, H_W(p), \lambda_{1,W}, \delta_W)]} \quad (13)$$

As shown in equation (7) ε is independent of p , therefore I can estimate equation (13) at the firm level:

$$1 - \frac{\gamma_W \sum_j N_j \times w^p(p_{j,t}, \beta_W, H_W(p), \lambda_{1,W}, \delta_W)}{\gamma_W \sum_j N_j \times w^p(p_{j,t}, \beta_M, H_W(p), \lambda_{1,W}, \delta_W)}$$

Where N_j is the number of female workers in each firm. In order to have a complete decomposition of the wage gap, I replace sequentially each female parameter for a male parameter until I reach the male predicted mean wage. Counterfactual wages are presented in Table 7

Differences in friction patterns imply differences in the observed distributions of firm productivity within the employed workers of different groups. This is true also when both groups face the same primitive distribution of productivity, see condition (6). Consequently, differences in frictions imply two effects over wages. The first one is the direct effect on the wage of a worker ε , who belong to the group k , working in a firm p , this is the effect

displayed on figure 1. And the second effect is the one that comes from aggregation due to changes in the distribution of accepted wages.

The wage setting equation implies that the higher the offer-arrival rate, the higher the wage, and the higher the job destruction rate the lower the wage, see Figure 1. But also increasing the offer-arrival rate makes the counterfactual firm's productivity distribution among workers to stochastically dominate the original one, while increasing the destruction rate has the opposite effect⁴⁷. Hence, the direct effect and the aggregation effect go in the same direction when we change λ_1 and δ .

I only observe the primitive distribution of productivity, that is the empirical distribution of p at the firm level, the distribution of productivity among males, and the distribution of productivity among females. I cannot estimate the wage equation with the counterfactual distribution of productivity only changing one friction parameter. Therefore, the counterfactual wages are calculated simulating the distribution of productivity faced by female workers using the friction parameters of male workers⁴⁸.

Considering the direct effect and the aggregation effect together the effect of frictions is significant. I find that on average across sectors, gender differences in destruction rates, explain 9% of the total wage gap, while decreasing

⁴⁷Given that the offer arrival rate of women is higher than the one of men, the distribution of firms productivity across the population of female worker stochastically dominates the counterfactual distribution of firm's productivity corresponding to female workers if they had the offer arrival rate of male workers. The proof is trivial, note that if $\lambda_1^F > \lambda_1^M$:

$$G(\lambda_1^W, \delta^W) = \frac{H(p)}{1 + \frac{\lambda_1^W}{\delta^W} \bar{H}(p)} \leq \frac{H(p)}{1 + \frac{\lambda_1^M}{\delta^W} \bar{H}(p)} = G(\lambda_1^M, \delta^W) \quad \blacksquare$$

On the other hand, given that women have higher destruction rates than men

$$G(\lambda_1^W, \delta^W) \leq G(\lambda_1^W, \delta^M).$$

⁴⁸The model used for simulations is a simplified version of the model presented in Section 2, where the worker heterogeneity has been omitted. Simulations use the punctual estimates of λ_1 , δ , γ_w , γ_u and α_l for every sector and worker group, reported in Section 3. I assume that the primitive distribution of firms productivity is log-normal. The mean of the distribution of firm's productivity is calibrated in equilibrium matching mean wages in each occupation group and in each sector.

MATA codes for simulating the model previously described are available from the author upon request.

Table 7: Counterfactual wages

COUNTERFACTUAL MEAN-DAILY WAGES		λ_1^M	λ_1^M	λ_1^M	λ_1^M	λ_1^M	λ_1^F
		δ^M	δ^M	δ^M	δ^M	δ^F	δ^F
		γ^M	γ^M	γ^M	γ^F	γ^F	γ^F
		$H(p)^M$	$H(p)^M$	$H(p)^F$	$H(p)^F$	$H(p)^F$	$H(p)^F$
		β^M	β^F	β^F	β^F	β^F	β^F
MANUFACT.	SK.	190.3	153.6	151.1	101.5	95.2	105.7
	UNSK.	109.2	114.1	116.2	78.1	71.0	76.9
CONSTRUCT.	SK.	150.9	103.1	104.6	73.3	69.5	82.8
	UNSK.	96.6	56.3	68.1	47.7	48.5	53.3
TRADE	SK.	119.2	84.7	92.8	74.6	69.7	67.8
	UNSK.	84.9	76.2	58.6	47.1	42.9	47.6
SERVICES	SK.	156.3	194.7	138.3	81.3	71.6	87.4
	UNSK.	88.5	97.6	80.6	47.4	45.7	49.1
WEIGHTED AVERAGE		124.1	113.1	103.9	70.0	65.1	72.1

the female offer arrival rate to be equal to the male one, would increase 13% the gap. Differences in δ are part of the gap, while differences in λ_1 , reduce the gap, therefore the effects of both type of frictions are partially compensated. The sizes of these effects are consistent with Bowlus (1997), where using samples of high school and college graduates from the National Longitudinal Survey of Youth (NLSY) these behavioral patterns were found to account for 20% - 30% of the wage differentials.

It is possible to disentangle the direct effect of frictions on wages from the aggregation effect. For this purpose I calculate the mean wage of women, changing the frictions parameters but keeping their original distribution of firms productivity as it is shown in Figure 1. These effects are surprisingly small, differences in δ explain, on average, 1.3 percent of the total wage gap, while netting out λ_1 would increase the wage-gap in 2.4 percent. In view of that, most of the effect of frictions comes through the aggregation effect.

The proportion of the wage gap that is due to differences in productivity connects directly with a branch of the literature initiated with Hellerstein et al (1999). In this line of work, they assumed equality between wages

and productivity, and therefore any inequality in wages that is not driven by differences in productivity may be considered as discrimination. Here wages and productivity are connected in a more sophisticated manner, in fact this relationship has been shown to be not an equality, not even for the non-discriminated group⁴⁹. On average, 65 percent of the total wage gap is accounted for by differences in productivity. The role of productivity in explaining the wage gap is large but not surprising given the definition of productivity used in this exercise. Behind these large productivity-gaps, there are important differences in productivity determinants. It can be seen on Table 2 that there are significant differences in age, tenure and potential experience. Moreover, there is also evidence of large differences in education attainment across gender in Germany⁵⁰.

There is a large literature studying the effect of segregation over the wage gap⁵¹. The structural estimation allows me to compare the current mean-wage of female workers with the counterfactual female mean-wage if they face the primitive distribution of firm's productivity faced by male workers (*ie*: changing $H_W(p)$ for $H_M(p)$). I find that 17 percent of the wage gap is accounted for by this difference. These results are consistent with results presented in Bayard, Hellerstein, Neumark and Troske (2003). Where, using matched employer-employee data from the U.S., they found a negative and small effect of the proportion of women in the establishment, over wages. One of the advantages of the estimation of this structural model is that it allows us to empirically disentangle the effect of segregation from differences in the distribution of firm's productivity generated by gender differences in friction patterns⁵².

The effect of differences in productivity is not stable across occupation, three quarters of wage gap is explained by differences in productivity in low qualification occupations, while this proportion reduces to 58% when con-

⁴⁹ $w_{ijt} = \varepsilon_i p_{jt} \Leftrightarrow \beta = 1$, but β is statistically different from one in every sector and in every worker group.

⁵⁰See Lauer (2003)

⁵¹See Altonji and Blank (1999).

⁵²To the best of my knowledge, up to the moment there is no paper disentangling both effects.

sidering high qualification occupations. Finally, these productivity measures are not on hourly-basis, and the numbers of hours significantly differ between male and female workers, see Section 5.1 for a more detailed discussion on this issue.

Table 8: Gender Wage-Gap decomposition

% OF THE WAGE-GAP EXPLAINED BY			
DIFFERENCES IN	UNSKILLED	SKILLED	ALL
δ	10.3	8.5	9.3
λ	-12.6	-13.8	-13.3
γ	74.5	58.3	65.1
$H(p)$	23.5	13.4	17.7
β	4.2	33.6	21.2
WAGE-GAP	39.0	44.3	41.9

On average, 21 percent of the total wage gap is accounted for by differences in the rent-splitting parameter. This means that women receive wages 9 percent lower than the ones received by equivalent men. This finding is different to what I obtained using the traditional approach based on Mincer-equations, where female workers are found to receive wages that are 15 percent lower than those of equivalent male workers (see Section A.6 in the appendix). As in the case of differences in productivity, differences in the wage setting parameter are also occupation-specific. As it can be seen in Table 8, this difference turns out to be an important determinant of the gender wage gap of workers in high qualification occupation, while it explains almost nothing of the wage gap in the low qualification jobs, this is consistent with the growing literature on glass ceiling.

5.1 Productivity-gap and differences in hours

One possible explanation to the large estimated productivity and wage gaps, may be that male workers work more hours than female ones. One of the main limitations of the LIAB is that it does not provide any measure of hours. Therefore the estimated differences in productivity, and the estimated dif-

ferences in wages are not on hourly-basis. In order to tackle this problem one alternative is to look for an external source of information about hours worked by each group in each sector. Using the GSOEP, I find significant differences in mean-hours between genders, see Table 9. On average, female workers are found to work almost 20% less hours than their male counterparts.

Table 9: Mean-Hours Per Week

		MEN	WOMEN	HOURS-GAP
MANUFACT.	UNSK.	39.95	38.84	12.8%
		(0.10)	(0.21)	(0.56%)
	SK.	41.68	38.69	7.2%
		(0.08)	(0.32)	(0.78%)
CONSTRUCT.	UNSK.	43.09	34.40	20.2%
		(0.26)	(1.65)	(3.85%)
	SK.	43.62	38.74	11.2%
		(0.12)	(0.97)	(2.24%)
TRADE	UNSK.	41.23	28.27	31.4%
		(0.47)	(0.45)	(1.35%)
	SK.	44.01	33.89	23.0%
		(0.24)	(0.52)	(1.25%)
SERVICES	UNSK.	44.67	28.12	37.0%
		(0.43)	(0.40)	(1.08%)
	SK.	46.33	42.08	9.2%
		(0.30)	(0.72)	(1.66%)

Note: Standard errors are given in parentheses. Means of weekly-hours are estimated using worker-level data from the GSOEP. Standard errors of the Hours-Gap are obtained by Delta-Method.

As the structural wage equation is linear in workers' ability, correcting for hours does not affect the estimated rent-splitting parameters⁵³. Nevertheless,

⁵³The hour-correction only modifies the worker productivity and the wage, if h is this correction:

$$hw_{i,j,t} = h\varepsilon_i w_{j,t,k(i)}^p(p_{j,t}, \beta_{k(i)}, H_{k(i)}(p), \lambda_{1k(i)}, \delta_{k(i)})$$

Therefore h is cancelled out of both sides of the equation and the original wage setting equation holds.

an interesting exercise is to have an hourly-wage-gap decomposition directly plugging the hour correction⁵⁴. Correcting mean-wages using mean-hours is going to be valid whenever hours worked are uncorrelated with wages, and there is evidence suggesting that this may be the case⁵⁵.

Table 10: Gender Wage-Gap decomposition

% OF THE HOURLY WAGE GAP EXPLAINED BY			
DIFFERENCES IN	UNSKILLED	SKILLED	ALL
δ	32.0	11.1	17.7
λ	-39.2	-18.1	-25.4
γ	43.8	21.8	48.3
$H(p)$	53.7	15.7	27.0
β	9.6	39.4	32.4
WAGE-GAP	17.1	37.8	27.5

Results are presented in Table 10. When correcting for hours, the wage gap is significantly smaller. I find that the average hourly-wage gap is 27.5%. A smaller fraction of this gap, 48 percent, is now explained by differences in productivity and on the other hand, a larger fraction of the gap, 32%⁵⁶, is due to differences in rent-splitting parameters. Segregation is responsible for 27 percent of the unconditional wage-gap in hourly basis. Netting out the offer arrival rate would increase the hourly-wage differential in 25 percent and if women had the male destruction rate, the gap would be 17 percent smaller.

6 Concluding Remarks

This paper presents the first estimation of an equilibrium search model using matched employer-employee data to study wage discrimination. This kind

⁵⁴I must thank Zvi Eckstein for this suggestion.

⁵⁵See Blundell and MaCurdy (1999).

⁵⁶As before female workers are receiving wages 6.2 percent lower than equivalent males due to differences in the rent-splitting parameter. This number does not change because the rent-splitting parameter estimates don't change.

of data is essential for structurally analyzing labor market discrimination mainly because:

- It allows us to have a clear distinction between differences in workers productivity across groups and differences in wage policies toward those groups.
- It makes the estimation simpler, and therefore more complete models may be estimated.
- It provides a clear way to empirically disentangle the effect of labor market segregation, from the effect of frictions on the distribution of accepted wages.

The structural estimation involved several steps: Firstly, I estimated group-specific productivity relying on production functions estimation at the firm-level using LIAB. Secondly, I computed job-retention and job-finding rates using GSOEP employee-level data. Finally, I estimated the wage-setting parameters (bargaining power) using individual wage records in LIAB, and transition parameters and productivity measures specific to each firm estimated in previous steps.

When analyzing productivity, I observe that women are 33 percent less productive than men in similar jobs, this difference is reduced to 17.5 percent if I control for differences in hours. The main findings in terms of friction patterns are that women are in general more mobile than men in terms of job-to-job transitions and they have also higher job-destruction rates. In spite of having large wage differentials, only in the construction sector women are found to have significantly lower rent-splitting parameters than men.

In terms of wages, I find that the unconditional gender wage gap is 42 percent. It turns out that most of the gap is accounted for by differences in productivity. Differences in destruction rates explain 9 percent of the total wage-gap and segregation is responsible for 17% of this wage differential. Netting out differences in offer-arrival rates would increase the gap by 13 percent. Differences in the rent-splitting parameter generate 21% of the

wage gap, which implies that female workers receive wages 9 percent lower than equivalent males.

There are two desirable extensions that I would like to perform. Firstly and more important, gender productivity gaps have been found to be very large, also controlling for differences in hours. I have proposed many robustness checks testing for different identification strategies, what ensures that the large gender productivity gap are not a statistical artifact. The main question, that stays unsolved after this paper, is understanding what is behind these gaps. Secondly, firm's productivity is estimated in a large N , but small T panel, hence there is an issue of a non-vanishing small- T measurement error in estimated firm's fixed effects. It would be interesting to obtain β estimates that take this problem into account.

A Appendix

A.1 Model Equations: Proofs

In this subsection, I derive analytically the close form solution of the equilibrium wage equation. The first step is to find the partial derivative with respect to the wage of the value of a job in a firm with productivity p for a worker with ability ε .

Applying the Leibniz integral rule in (1).

$$\frac{\partial [E(w(p, \varepsilon), \varepsilon)]}{\partial w(p, \varepsilon)} = \frac{1}{(r + \delta + \lambda_1 \bar{F}(w(p, \varepsilon) | \varepsilon))} \quad (14)$$

Integrating (14) between $w(p_{\min}, \varepsilon)$ and $w(p, \varepsilon)$.

$$\begin{aligned} \int_{w(p_{\min}, \varepsilon)}^{w(p, \varepsilon)} \frac{1}{(r + \delta + \lambda_1 \bar{F}(\tilde{w}(p, \varepsilon) | \varepsilon))} d(\tilde{w}(p, \varepsilon)) &= \int_{w(p_{\min}, \varepsilon)}^{w(p, \varepsilon)} \frac{\partial [E(\tilde{w}(p, \varepsilon), \varepsilon)]}{\partial \tilde{w}(p, \varepsilon)} d(\tilde{w}(p, \varepsilon)) \\ E(w(p, \varepsilon), \varepsilon) - E(w(p_{\min}, \varepsilon), \varepsilon) &= E(w(p, \varepsilon), \varepsilon) - U(\varepsilon) \end{aligned}$$

Using the surplus splitting rule (3), the value of the job for the worker (1), the value of the job for the firm (2) and rearranging:

$$\begin{aligned}
w(p, \varepsilon) &= p\varepsilon - \\
&(\rho + \delta + \lambda_1 \bar{F}(w(p, \varepsilon)|\varepsilon)) \frac{(1 - \beta)}{\beta} \\
&* \int_{w(p_{\min}, \varepsilon)}^{w(p, \varepsilon)} \frac{1}{(\rho + \delta + \lambda \bar{F}(\tilde{w}(p, \varepsilon)|\varepsilon))} d(\tilde{w}(p, \varepsilon))
\end{aligned} \tag{15}$$

noting that

$$\begin{aligned}
&\int_{w(p_{\min}, \varepsilon)}^{w(p, \varepsilon)} \frac{1}{(\rho + \delta + \lambda \bar{F}(\tilde{w}(p, \varepsilon)|\varepsilon))} d(\tilde{w}(p, \varepsilon)) \\
&= \int_{p_{\min}}^p \frac{1}{(\rho + \delta + \lambda \bar{H}(p'))} \frac{d(w(p, \varepsilon))}{dp'} dp'
\end{aligned}$$

and taking derivatives with respect to p .

$$\begin{aligned}
\frac{d(w(p, \varepsilon))}{dp'} &= \varepsilon - \frac{(1 - \beta)}{\beta} \frac{d(w(p, \varepsilon))}{dp'} \\
&+ \lambda_1 h(p) \frac{(1 - \beta)}{\beta} \int_{p_{\min}}^p \frac{1}{(\rho + \delta + \lambda \bar{H}(p'))} \frac{d(w(p, \varepsilon))}{dp'} dp'
\end{aligned}$$

Then, plugging equation (15):

$$\begin{aligned}
\frac{d(w(p, \varepsilon))}{dp'} &= \varepsilon + \\
&\lambda_1 h(p) \frac{w(p, \varepsilon) - p\varepsilon}{(\rho + \delta + \lambda \bar{H}(p'))} - \frac{(1 - \beta)}{\beta} \frac{d(w(p, \varepsilon))}{dp'}
\end{aligned}$$

Rearranging, I have a first order differential equation,

$$\frac{d(w(p, \varepsilon))}{dp'} + \frac{\beta \lambda_1 h(p)}{\rho + \delta + \lambda_1 \bar{H}(p)} w(p, \varepsilon) = \varepsilon \beta \left[\frac{\rho + \delta + \lambda_1 \bar{H}(p) + \lambda_1 h(p)p}{\rho + \delta + \lambda_1 \bar{H}(p)} \right] \tag{16}$$

To solve this differential equation, note that:

$$\frac{d(\rho + \delta + \lambda_1 \bar{H}(p))^{-\beta}}{dp} = (\rho + \delta + \lambda_1 \bar{H}(p))^{-\beta} \frac{\beta \lambda_1 h(p)}{\rho + \delta + \lambda_1 \bar{H}(p)}$$

Then, multiplying both sides of equation (16) by $(\rho + \delta + \lambda_1 \bar{H}(p))^{-\beta}$ and rearranging

$$\frac{d [w(p, \varepsilon)(\rho + \delta + \lambda_1 \bar{H}(p))^{-\beta}]}{dp} = \varepsilon \beta \left[\frac{\rho + \delta + \lambda_1 \bar{H}(p) + \lambda_1 h(p)p}{(\rho + \delta + \lambda_1 \bar{H}(p))^{1+\beta}} \right] \quad (17)$$

Integrating (17) between p_{\min} and p , and noting that the lowest productivity firm will produce no surplus $\Leftrightarrow w(p_{\min}, \varepsilon) = p_{\min} \varepsilon$, straightforward algebra shows that:

$$\begin{aligned} & w(p, \varepsilon)(\rho + \delta + \lambda_1 \bar{H}(p))^{-\beta} \\ &= (\rho + \delta + \lambda_1)^{-\beta} p_{\min} \varepsilon + \varepsilon \beta \int_{p_{\min}}^p \left[\frac{\rho + \delta + \lambda_1 \bar{H}(p') + \lambda_1 h(p')p'}{(\rho + \delta + \lambda_1 \bar{H}(p'))^{1+\beta}} \right] dp' \end{aligned}$$

separating the integral in a convenient way and noting that:

$$\frac{\partial \left((\rho + \delta + \lambda_1 \bar{H}(p'))^{-\beta} p' \right)}{\partial p'} = (\rho + \delta + \lambda_1 \bar{H}(p'))^{-\beta} + \frac{\beta \lambda_1 h(p') p'}{(\rho + \delta + \lambda_1 \bar{H}(p'))^{1+\beta}} dp'$$

it solves as:

$$\begin{aligned} w(p, \varepsilon) &= \frac{(\rho + \delta + \lambda_1 \bar{H}(p))^\beta}{(\rho + \delta + \lambda_1)^\beta} p_{\min} \varepsilon - \\ &\quad \varepsilon (1 - \beta) (\rho + \delta + \lambda_1 \bar{H}(p))^\beta \int_{p_{\min}}^p (\rho + \delta + \lambda_1 \bar{H}(p'))^{-\beta} dp' + \\ &\quad \varepsilon (\rho + \delta + \lambda_1 \bar{H}(p))^\beta \int_{p_{\min}}^p \frac{\partial \left((\rho + \delta + \lambda_1 \bar{H}(p'))^{-\beta} p' \right)}{\partial p'} dp' \end{aligned}$$

rearranging I get the wage equation as a function of individual skill (ε), friction patterns (δ and λ_1) and firm's productivity (p).

$$w(p, \varepsilon) = \varepsilon p - \varepsilon (1 - \beta) (\rho + \delta + \lambda_1 \bar{H}(p))^\beta \int_{p_{\min}}^p (\rho + \delta + \lambda_1 \bar{H}(p'))^{-\beta} dp' \quad \blacksquare$$

Now I show that p_{\min} is independent of ε . p_{\min} is the minimum observed productivity level. Firms with productivity p_{\min} make zero profit, and therefore the whole productivity goes to the worker, who receive εp_{\min} this wage exactly compensate the worker to leave the unemployment, Therefore:

$$E(p_{\min}\varepsilon, \varepsilon) = U(\varepsilon)$$

$$\begin{aligned} & p_{\min}\varepsilon + \lambda_1 \int_{w(p_{\min}, \varepsilon)}^{w(p_{\max}, \varepsilon)} [E(w(p', \varepsilon), \varepsilon) - U(\varepsilon)] dF(W(p', \varepsilon)) \\ = & b\varepsilon + \lambda_0 \int_{w(p_{\min}, \varepsilon)}^{w(p_{\max}, \varepsilon)} [E(w(p', \varepsilon), \varepsilon) - U(\varepsilon)] dF(W(p', \varepsilon)) \end{aligned}$$

$$p_{\min}\varepsilon = b\varepsilon + (\lambda_0 - \lambda_1) \int_{w(p_{\min}, \varepsilon)}^{w(p_{\max}, \varepsilon)} [E(w(p', \varepsilon), \varepsilon) - U(\varepsilon)] dF(W(p', \varepsilon))$$

Using the surplus splitting rule (3):

$$p_{\min}\varepsilon = b\varepsilon + (\lambda_0 - \lambda_1) \frac{\beta}{1 - \beta} \int_{w(p_{\min}, \varepsilon)}^{w(p_{\max}, \varepsilon)} \frac{p'\varepsilon - w(p', \varepsilon)}{(\rho + \delta + \lambda \bar{F}(w(p', \varepsilon)|\varepsilon))} dF(W(p', \varepsilon))$$

This is the value function for a worker of a given ε , therefore we can rearrange everything in terms of p .

$$p_{\min}\varepsilon = b\varepsilon + (\lambda_0 - \lambda_1) \frac{\beta}{1 - \beta} \int_{p_{\min}}^{p_{\max}} \frac{p'\varepsilon - w(p', \varepsilon)}{(\rho + \delta + \lambda \bar{H}(p'))} dH(p')$$

using equation (4) and rearranging:

$$\begin{aligned} p_{\min}\varepsilon &= b\varepsilon + \varepsilon(\lambda_0 - \lambda_1) \frac{\beta}{1 - \beta} \\ &\quad \times \int_{p_{\min}}^{p_{\max}} \frac{(1 - \beta) \int_{p_{\min}}^{p'} (\rho + \delta + \lambda_1 \bar{H}(\tilde{p}))^{-\beta} d\tilde{p}}{(\rho + \delta + \lambda \bar{H}(p'))^{(1-\beta)}} dH(p') \end{aligned}$$

ε becomes irrelevant:

$$p_{\min} = b + (\lambda_0 - \lambda_1) \beta \int_{p_{\min}}^{p_{\max}} \frac{\int_{p_{\min}}^{p'} (\rho + \delta + \lambda_1 \bar{H}(\tilde{p}))^{-\beta} d\tilde{p}}{(\rho + \delta + \lambda \bar{H}(p'))^{(1-\beta)}} dH(p') \quad \blacksquare$$

Note that p_{\min} is a function of the distribution of p and the parameters of the model. The intuition, in discrete time, is clear because the value of being employed and the value of being unemployed are infinite additions of flows which are linear on ε ($w(\varepsilon, p)$ and $b\varepsilon$). Each flow is multiplied by the discount rate and the probability of being in each state, that do not depend on ε . Hence the value of being employed and the value of being unemployed are both linear in ε . This condition must hold in order to avoid sorting between p and ε .

A.2 Duration model - Maximum Likelihood Specification

The unconditional likelihood of job spell durations is:

$$\mathcal{L}(t) = \int \mathcal{L}(t|p)g(p)dp.$$

$$\mathcal{L}(t) = \int_{p_{\min}}^{p_{\max}} \frac{(1 + \kappa_1)h(p)}{1 + \kappa_1\bar{H}(p)} [\delta + \lambda_1 H(p)] e^{-[\delta + \lambda_1 H(p)]t} dp$$

Rearranging and recalling that $\kappa_1 = \lambda_1/\delta$.

$$\mathcal{L}(t) = \frac{(1 + \kappa_1)\delta}{\kappa_1} \int_{p_{\min}}^{p_{\max}} \frac{1}{\delta + \lambda_1\bar{H}(p)} e^{-[\delta + \lambda_1 H(p)]t} \lambda_1 h(p) dp$$

Changing the variable within the integral, $x = [\delta + \lambda_1\bar{H}(p)]t$. After straightforward algebra I get:

$$\mathcal{L}(t) = \frac{(1 + \kappa_1)\delta}{\kappa_1} [E_1(\delta t) - E_1(\delta(1 + \kappa_1)t)]$$

where $E_1(t) = \int_t^\infty \frac{e^{-x}}{x} dx$ is the exponential integral function ■.

My sample covers a fixed number of periods, so that some job durations are right censored, and other job spells started before the panel's beginning. Then, the exact likelihood function that takes into account these events is:

$$l(t_i) = (1 - c_i) \log \left(\frac{\mathcal{L}(t_i)}{\int_{H_i}^\infty \mathcal{L}(t) dt} \right) + c_i \log \left(\frac{\int_{t_i}^\infty \mathcal{L}(t) dt}{\int_{H_i}^\infty \mathcal{L}(t) dt} \right)$$

where c_i is a truncated spell indicator and H_i is the time period elapsed before the sample.

$$l(t_i) = (1 - c_i) \log \left(\frac{[E_1(\delta t) - E_1(\delta(1 + \kappa_1)t)]}{\int_{H_i}^{\infty} [E_1(\delta t) - E_1(\delta(1 + \kappa_1)t)] dt} \right) \\ + c_i \log \left(\frac{\int_{t_i}^{\infty} [E_1(\delta t) - E_1(\delta(1 + \kappa_1)t)] dt}{\int_{H_i}^{\infty} [E_1(\delta t) - E_1(\delta(1 + \kappa_1)t)] dt} \right)$$

Using the fact that $\int E_1(at)dt = -\int E_i(-at)dt = -\left(tE_i(-at) + \frac{e^{-at}}{a}\right)$ (see Abramowitz and Stegun, 1972), and noting that $E_1(-\infty) = 0$

$$\int_{t_i}^{\infty} [E_1(\delta t) - E_1(\delta(1 + \kappa_1)t)] dt = \int_{t_i}^{\infty} E_1(\delta t)dt - \int_{t_i}^{\infty} E_1(\delta(1 + \kappa_1)t)dt \\ = -tE_i(-\delta t) + \frac{e^{-\delta t}}{\delta} \Big|_{t_i}^{\infty} + \\ tE_i(-(1 + \kappa_1)\delta t) + \frac{e^{-\delta t}}{(1 + \kappa_1)\delta} \Big|_{t_i}^{\infty}$$

$$\int_{t_i}^{\infty} [E_1(\delta t) - E_1(\delta(1 + \kappa_1)t)] dt \\ = t_i E_i(-\delta t_i) + \frac{e^{-\delta t_i}}{\delta} - t_i E_i(-(1 + \kappa_1)\delta t) - \frac{e^{-\delta t_i}}{(1 + \kappa_1)\delta}$$

since $E_i(-at) = -E_1(at) = -\int_{at_i}^{\infty} \frac{e^{-x}}{x} dx$.

$$\int_{t_i}^{\infty} [E_1(\delta t) - E_1(\delta(1 + \kappa_1)t)] dt \\ = \frac{e^{-\delta t_i}}{\delta} - t_i \int_{\delta t_i}^{\delta(1 + \kappa_1)t} \frac{e^{-x}}{x} dx - \frac{e^{-\delta t_i}}{(1 + \kappa_1)\delta}$$

The same is true for $\int_{H_i}^{\infty} \mathcal{L}(t)dt$. Then the likelihood takes the following form:

$$\begin{aligned}
l(t_i) = & (1 - c_i) \log \left(\frac{\int_{\delta t}^{(1+\kappa_1)\delta t} \frac{e^{-x}}{x} dx}{\frac{e^{-\delta H_i}}{\delta} - \frac{e^{-\delta(1+\kappa_1)H_i}}{\delta(1+\kappa_1)} - H_i \int_{\delta H_i}^{(1+\kappa_1)\delta H_i} \frac{e^{-x}}{x} dx} \right) + \\
& c_i \log \left(\frac{\frac{e^{-\delta t_i}}{\delta} - \frac{e^{-\delta(1+\kappa_1)t_i}}{\delta(1+\kappa_1)} - t_i \int_{\delta t_i}^{(1+\kappa_1)\delta t_i} \frac{e^{-x}}{x} dx}{\frac{e^{-\delta H_i}}{\delta} - \frac{e^{-\delta(1+\kappa_1)H_i}}{\delta(1+\kappa_1)} - H_i \int_{\delta H_i}^{(1+\kappa_1)\delta H_i} \frac{e^{-x}}{x} dx} \right) \quad \blacksquare
\end{aligned}$$

A.3 Robustness Check 1: Production Function

The robustness of the productivity estimates are crucial to be able to reach reliable conclusions about wage discrimination. As robustness a check, I present results of the productivity estimates under different sets of assumptions.

No Frictions in the Capital Market: As it has been assumed in the theory, the labor input is given for the firm because it has not control over job-creation and job-destruction Poisson processes but capital is chosen to maximize profit. Assuming that there are no frictions or adjustment cost in the capital market, when a firm knows the total labor L_{jt} it will have in the current period, it solves the following problem:

$$\max_{k_{jt}} (A_j K_{jt}^{(1-\alpha)} L_{jt}^\alpha - r_t K_{jt})$$

Substituting the first order condition into the production function and rearranging, I have that:

$$Y_{jt} = \left[\left(A_j^{\frac{1}{1-\alpha}} \frac{1-\alpha}{r_t} \right)^{\frac{1-\alpha}{\alpha}} e^{u_{jt}} \right] L_{jt} = p_{jt} L_{jt}$$

where r_t is the cost of capital. Note that this production function is equivalent to $p\varepsilon$, the production function assumed in the theory, where p is time and firm-specific: $p_{jt} = A_j^{\frac{1}{\alpha}} \left(\frac{1-\alpha}{r_t} \right)^{\frac{1-\alpha}{\alpha}} e^{u_{jt}}$.

$$\log(Y_{jt}) = a_j + b_t + \log(L_{jt}^{ms} + \gamma_w L_{jt}^{ws} + \gamma_u L_{jt}^{mu} + \gamma_w \gamma_u L_{jt}^{wu}) + u_{jt} \quad (18)$$

In Table 11, I report within-groups non linear least squares estimates of (8) and (18). Relative productivity estimates (ie: γ_w and γ_u) are very similar in both estimations, differences between parameters are always smaller than a standard deviation. This robustness check is essential because in this kind of model, it is assumed that capital adjusts instantaneously to match the number of workers in each period. Although this assumption may seem controversial, it turns out that using observed capital or using the theoretical optimal choice of capital does not change the relative productivity estimates.

Table 11: Production Function: Optimal Capital Input

	WG-NLLS OF (8)		WG-NLLS OF (18)	
	γ_w	γ_u	γ_w	γ_u
MANUFACTURING	0.672 (0.062)	0.484 (0.042)	0.642 (0.132)	0.478 (0.070)
CONSTRUCTION	0.701 (0.052)	0.444 (0.039)	0.839 (0.141)	0.485 (0.833)
TRADE	0.804 (0.092)	0.487 (0.056)	0.745 (0.176)	0.541 (0.982)
SERVICES	0.588 (0.068)	0.298 (0.030)	0.595 (0.115)	0.301 (0.040)

Note: Time dummies included. Robust to heteroskedasticity Standard errors are given in parentheses.

Worker Composition Endogeneity: One of the main criticisms to the productivity estimates reported in Hellerstein and Neumark (1999) and in Hellerstein, Neumark, and Troske (1999) was that the proportion of women in the firm is likely to be correlated with the firm's technology⁵⁷.

$$\begin{aligned}
\log(Y_{jt}) = & \log(\check{A}_j) + \alpha_k \log(K_{jt}^d) + \\
& \alpha_l \log(L_{jt}^{ms} + \gamma_w L_{jt}^{ws} + \gamma_u L_{jt}^{mu} + \gamma_w \gamma_u L_{jt}^{wu}) + u_{jt}
\end{aligned} \tag{19}$$

Note that (19) is the original Cobb-Douglas production function in logs without imposing constant returns to scale and including the depreciated

⁵⁷See Altonji and Blank (1999).

Table 12: Production Function: Non Linear Least Squares in Levels

NLLS OF (19)				
	α_k	α_l	γ_w	γ_u
MANUFACTURING	0.153 (0.007)	0.905 (0.018)	0.352 (0.032)	0.300 (0.016)
CONSTRUCTION	0.084 (0.017)	1.034 (0.028)	0.451 (0.082)	0.382 (0.043)
TRADE	0.110 (0.018)	0.963 (0.029)	0.562 (0.068)	0.365 (0.039)
SERVICES	0.180 (0.098)	0.839 (0.015)	0.356 (0.029)	0.292 (0.017)

Note: Time dummies included. Standard errors are given in parentheses.

capital instead of the optimal capital input. In this estimation I am assuming that the depreciation rate is constant, and hence depreciated capital is a constant fraction of the total capital.

The results are presented in Table 12. Female productivity estimates are significantly smaller. This finding is true for all sectors. Comparing these results with results presented in Table (4), where firm's fixed effects were removed, confirms the Altonji & Blank (1999) suspicion about correlation between the women proportion and the firm fixed effect. Hausman tests reject equality of γ_w in every sector. In terms of γ_u , the results are not so different and I only reject equality for manufacturing.

Predetermined inputs:

Estimating (8) the firm fixed effect is completely removed, but the simultaneity problem is not totally solved. One alternative would be to treat L_{jt} and k_{jt} as predetermined variables. In Table 13 I report System-GMM estimates of (19). However the precision in the γ 's GMM estimates is poor. Capital coefficients are not significantly different from zero and constant returns to scale are not rejected in any sector. Sargan tests do not reject compatibility of instruments in any sector. Hausman test of equality between γ 's reported in Table 13 and those in Table 4 do not reject equality in any group.

Table 13: Production Function: Non Linear SYSTEM-GMM

SYSTEM-GMM OF (19)					
	α_k	α_l	γ_w	γ_u	SARGAN P-V
MANUFACTURING	0.060 (0.046)	0.938 (0.011)	0.70 (0.454)	0.15 (0.069)	58%
CONSTRUCTION	0.016 (0.043)	1.164 (0.160)	0.442 (0.267)	0.188 (0.083)	98%
TRADE	0.053 (0.050)	1.003 (0.170)	0.370 (0.290)	0.322 (0.215)	97%
SERVICES	0.117 (0.071)	0.786 (0.146)	0.434 (0.158)	0.204 (0.090)	93%

Note: Time dummies included. Standard errors are given in parentheses.

A.4 Robustness Check 2: Constructing firm's productivity

The productivity of firm j in time t , is constructed with the parameters obtained in Section 4.1. Although the production function was estimated using output, productivity is constructed using value-added which is conceptually more accurate.

There are two possibilities to construct marginal productivity measures:

- To include the residual term in the production function:

$$\check{p}_{jt} = \alpha_l A_j K_{j,t}^{(1-\alpha_l)} (L_{jt})^{(\alpha_l-1)} e^{u_{jt}} = \alpha_l \frac{Y_{jt}}{L_{jt}}$$

- To ignore this residual term:

$$\tilde{p}_{jt} = \alpha_l A_j K_{j,t}^{(1-\alpha_l)} (L_{jt})^{(\alpha_l-1)}$$

For simplicity and following Cahuc et al (2006), I choose the first option. In this case the productivity is simply $p_{jt} = \alpha_l Y_{jt}/L_{jt}$. Whether firms insure temporary shocks to workers, is an open debate in the literature (see Guiso,

Table 14: Robustness Check: Productivity

		WOMEN		MEN	
		WITH $e^{u_{jt}}$	WITHOUT $e^{u_{jt}}$	WITH $e^{u_{jt}}$	WITHOUT $e^{u_{jt}}$
		β	β	β	β
M	UNSK.	0.419	0.438	0.398	0.551
		(0.129)		(0.096)	
	SK.	0.226	0.134	0.292	0.245
		(0.088)		(0.046)	
C	UNSK.	0.214	0.195	0.408	0.407
		(0.090)		(0.045)	
	SK.	0.113	0.109	0.186	0.286
		(0.078)		(0.104)	
T	UNSK.	0.339	0.326	0.382	0.435
		(0.173)		(0.106)	
	SK.	0.152	0.135	0.222	0.194
		(0.066)		(0.064)	
S	UNSK.	0.849	0.794	0.757	0.686
		(0.125)		(0.125)	
	SK.	0.413	0.423	0.324	0.402
		(0.104)		(0.073)	

Note: Bootstrap Standard errors in parentheses for the model with $e^{u_{jt}}$

Pistaferri and Schivardi, 2005). If firms totally insure temporary shocks to workers, and wages are determined as a function of the expected marginal productivity, \tilde{p}_{jt} would be the relevant measure. If not, \check{p}_{jt} would be the correct one. In Table 14 I show that estimated β with or without including the residual term are very similar. Results with $e^{u_{jt}}$ (*ie.* using \check{p}_{jt}) are the same estimates reported in Table 6

A.5 Robustness Check 3: Allowing for Between Firms Bertrand Competition

In the model presented in Section 2, workers do not have the option of recalling old employers. In this Subsection I estimate the model allowing recalling and Bertrand competition between firms as in Cahuc, Postel-Vinay

Table 15: Robustness Check: Allowing for Renegotiation

		WOMEN	MEN
		β_{CPR}	β_{CPR}
MANUFACTURING	UNSKILLED	0.212	0.182
	SKILLED	0.163	0.172
CONSTRUCTION	UNSKILLED	0.223	0.289
	SKILLED	0.182	0.206
TRADE	UNSKILLED	0.238	0.254
	SKILLED	0.203	0.215
SERVICES	UNSKILLED	0.325	0.341
	SKILLED	0.241	0.231

Note: β_{CPR} is the Nash bargaining power of the worker in the model with renegotiation proposed in Cahuc, Postel-Vinay and Robin (2006). β_{CPR} are recovered by simulated method of moments.

and Robin (2006). β has a different interpretation in this model, it is still a surplus-splitting parameter where the surplus has been defined in terms of a time varying outside option given by a poaching firm⁵⁸.

The estimated bargaining power are smaller than in the model without Bertrand competition, now the weighted average is 21.8%. I find similar patterns in terms of gender, than in the model without renegotiation. Women are found to have smaller bargaining power than men in most of the groups. As in the model proposed in this paper, female workers are only found to have larger β in services and in manufacturing but only in low qualification occupations.

Workers in low qualification occupation are found to have higher bargaining power than workers in high qualification occupation. This results have been found also estimating the model without renegotiation but it is different to what has been found by Cahuc, Postel-Vinay and Robin (2006), estimating a similar model with French data.

The counterfactual decomposition works in the same way as the decomposition described in Section 5. I first calculate the mean wages of female

⁵⁸For the exact formulation of the bargaining scenario and a discussion on its implication see Cahuc, Postel-Vinay and Robin (2006).

workers, as a function of female wage determinants, and I sequentially change each parameter until reaching the male mean wages.

Table 16: Gender Wage-Gap decomposition

MODEL WITH BERTRAND COMPETITION			
DIFFERENCES IN	UNSKILLED	SKILLED	ALL
δ	19.9	19.5	19.7
λ	-12.7	-7.1	-9.4
γ	91.9	77.9	83.5
β	0.1	9.7	6.2

The decomposition is similar to the one that comes out from the model without Bertrand Competition. Now 6.2 % of the wage gap is explained by differences in the bargaining power, slightly less than before. Female workers in high qualification occupations are suffering more wage discrimination. Differences in productivity are responsible for most of the wage gap. Allowing for Bertrand competition also increases the effect of differences in destruction rates, decreases the effect of differences in job-offers arrival rates, and the net effect of friction is now more important.

Details of the simulations:

The model used for simulations is a simplified version of the Cahuc, Postel-Vinay and Robin (2006) model, where the worker heterogeneity has been omitted⁵⁹.

- β_{CPR} are recovered by the simulated method of moments.
- Simulations use the punctual estimates of λ_1 , δ , γ_w , γ_u and α_l for every sector and worker group, reported in Section 3.
- I assume that the primitive distribution of firm's productivity is log-normal.

⁵⁹Given that I am matching means, and the wage equation is linear in worker ability, worker heterogeneity does not play any role in these simulations.

MATA codes for simulating the model previously described are available from the author upon request.

- 32 moments have been matched
 - The mean-wages of female and male workers in each occupation group and in each sector.
 - The mean-productivity of the endogenously truncated distribution of firms faced by female and male workers in each occupation group and in each sector.
- Using condition (5), the unemployment rate of each group reported in EUROSTAT⁶⁰ and the estimates of δ for each group, I recover an estimate of λ_0 for female and male workers in each occupation group and in each sector.

A.6 Detecting Discrimination - Traditional Approach

In order to compare different strategies to detect wage discrimination. I perform the traditional approach using Mincer-type wage equations. As it can be seen in Table 17, women have large wage differentials. Controlling for observable characteristics, they receive wages, on average, 21 percent lower than men. This difference is significant and consistent with what has been found in previous research: Blau and Kahn (2000), with OECD data reports a difference of 25.5 percent between male and female mean wages, while Fitzenberger and Wunderlich (2002) with the same data as in this paper, but using quantile regression, the estimated German gender wage gap ranges between 16 percent and 25 percent depending on job's qualification.

Oaxaca-Blinder Decomposition

Using the results presented in Table 17, I calculate a Oaxaca-Blinder decomposition, which is basically to decompose the wage-gap between differences in observable and unobservable characteristics. The counterfactual female mean-wage has to be interpreted as the mean-wage that women would have if they had the male distribution of observable characteristics. Therefore, the difference between the counterfactual female mean-wage and the

⁶⁰The mean unemployment rate between 1996 and 2005 was 9.64 percent for females and 9.11 percent for males (see EUROSTAT).

Table 17: Mincer Wage Equations - Censored-Normal Regression. Maximum Likelihood Estimates

Y=LOG(WAGE)	ALL	MEN	WOMEN
WOMEN	-0.211 (0.0004)	- -	- -
IMMIGRANT	0.073 (0.0006)	0.061 (0.0016)	0.076 (0.0006)
SKILLED	0.255 (0.0004)	0.178 (0.0010)	0.276 (0.0005)
AGE	0.056 (0.0002)	0.068 (0.0004)	0.054 (0.0002)
AGE ²	-0.001 (0.0000)	-0.001 (0.0000)	-0.001 (0.0000)
PRIMARY EDUCATION	0.236 (0.0004)	0.257 (0.0011)	0.234 (0.0005)
COLLEGE (INCOMPLETE)	-0.127 (0.0014)	-0.082 (0.0028)	-0.162 (0.0015)
TECHNICAL COLLEGE (COMPLETED)	0.386 (0.0010)	0.436 (0.0021)	0.354 (0.0012)
COLLEGE	0.609 (0.0011)	0.616 (0.0033)	0.566 (0.0011)
UNIVERSITY DEGREE	0.757 (0.0011)	0.819 (0.0027)	0.700 (0.0012)
TENURE	0.017 (0.0001)	0.025 (0.0002)	0.015 (0.0001)
EXPERIENCE	0.033 (0.0001)	0.021 (0.0003)	0.036 (0.0001)
PART-TIME	-0.638 (0.0007)	-0.651 (0.0010)	-0.608 (0.0011)
MANUFACTURING	0.178 (0.0008)	0.175 (0.0016)	0.103 (0.0010)
CONSTRUCTION	0.063 (0.0013)	0.026 (0.0029)	-0.081 (0.0014)
SERVICES	0.037 (0.0009)	0.025 (0.0017)	-0.023 (0.0011)
CONSTANT	2.500 (0.0029)	2.189 (0.0066)	2.599 (0.0031)
PSEUDO R ²	47.23	30.92	52.53
SIGMA	0.38	0.48	0.34

Note: Std. errors are given in parentheses. Time Dummies included.

observed women mean-wage is the portion of the gap understood as discrimination.

Following this approach, I would conclude that women are being discriminated. They are receiving wages which are on average almost 15 percent lower than wages of similar men in terms of observable characteristics. These results are slightly different to those obtained in this paper what might be suggesting that, in this case, the traditional approach is not able to disentangle differences in unobserved characteristics from discrimination.

References

- [1] Abowd, J. M., Kramarz, F., and Margolis, D. N. (1999), "High wage workers and high wage firms". *Econometrica*, 67 (2), 251-334.
- [2] Abramowitz, M., and A. Stegun (1972), "Handbook of Mathematical Functions, with Formulas, Graphs, and Mathematical Tables" (Tenth Ed.), National Bureau of Standards Applied Mathematics Series, Vol. 55. Washington, DC: U.S. Government Printing Office.
- [3] Alda H., St. Bender, H. Gartner (2005), "The linked employer-employee dataset of the IAB (LIAB)" IAB-Discussion Paper.6/2005.
- [4] Altonji, J.G. and R.M. Blank (1999), "Race and Gender in the Labor Market", in: O. Ashenfelter and D. Card (eds.), *Handbook of Labor Economics*, Volume 3, Amsterdam: Elsevier Science.
- [5] Arellano, M. and S.R. Bond (1991), "Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations", *Review of Economic Studies*, 58, 277–297.
- [6] Arellano, M., and O. Bover (1995), "Another Look at the Instrumental-Variable Estimation of Error-Components Models," *Journal of Econometrics*, 68, 29–52.
- [7] Atakan, A. (2006) "Assortative matching with explicit search cost" *Econometrica*, 74 (3), 667-680
- [8] Bayard, K., J. Hellerstein, D. Neumark, and K. Troske. 2003. "New Evidence on Sex Segregation and Sex Differences in Wages from Matched Employee-Employer Data." *Journal of Labor Economics* 21 (4), 887-922.
- [9] Becker, G. (1973), "A Theory of Marriage: Part I," *Journal of Political Economy*, 81, 813-846.
- [10] Blau. F. and L. Kahn (2003), "Understanding International Differences in the Gender Pay Gap", *Journal of Labor Economics*, 21(1): 106-144.

- [11] Blau. F. and L. Kahn (2000), "Gender Differences in Pay" *Journal of Economic Perspectives* 14(49) 75–99
- [12] Blundell R. and S. Bond (1999), "GMM Estimation with Persistent Panel Data: An Application to Production Functions", IFS WP n° W99/4.
- [13] Blundell R and T. MaCurdy (1999) "Labor supply: A review of alternative approaches" in: O. Ashenfelter and D. Card (eds.), *Handbook of Labor Economics*, Volume 3, Amsterdam: Elsevier Science.
- [14] Bontemps, C., J.-M. Robin, and G. J. van den Berg (2000), "Equilibrium Search with Continuous Productivity Dispersion: Theory and Non-Parametric Estimation," *International Economic Review*, 41, 305–358.
- [15] Bowlus, A. (1997), "A Search Interpretation of Male-Female Wage Differentials," *Journal of Labor Economics*, 15(4), 625-657.
- [16] Bowlus, A. and Z. Eckstein (2002), "Discrimination and Skill Differences in an Equilibrium Search Model," *International Economic Review*, 43(4), 1309-1345.
- [17] Burdett, K. and D. T. Mortensen (1998), "Wage-Differentials, Employer Size and Unemployment, *International Economic Review*, 39, 257–273.
- [18] Cahuc, P., F. Postel-Vinay and J.-M. Robin, (2006), "Wage Bargaining with On-the-job search: Theory and Evidence", *Econometrica*, 74(2), 323-64.
- [19] Chernozhukov, V. and H. Hong, (2003), "An MCMC Approach to Classical Estimation" *Journal of Econometrics* 115 (2), 293-346
- [20] Crepon B. ,N. Deniau and S. Pérez-Duarte (2002), "Wages, Productivity, and Worker Characteristics: A French Perspective" CREST WP 2003/04

- [21] Croson, R., U. Gneezy, "Gender Differences in Preferences" *Journal of Economic Literature*, Volume 47, Number 2, June 2009 , pp. 448-474(27).
- [22] Dong X, and L Zhang (2009) "Economic transition and gender differentials in wages and productivity: Evidence from Chinese manufacturing enterprises" *Journal of Development Economics* 88(1), pp. 144-156
- [23] Eckstein, Z. and K.I.Wolpin (1995), "Duration to first job and return to schooling: estimates from a search-matching model," *Review of Economic Studies*, 62, 263-286.
- [24] Eckstein, Z. and K. Wolpin (1999), "Estimating the Effect of Racial Discrimination on First Job Wage Offers", *The Review of Economics Statistics*, 81(3): 384-392.
- [25] Flabbi L.(2010), "Gender Discrimination Estimation in a Search Model with Matching and Bargaining" *International Economic Review*, 51(3): 745-783.
- [26] Flinn, C. (2006), "Minimum Wage Effects on Labor Market Outcomes under Search, Bargaining, and Endogenous Contact Rates." *Econometrica* 74 (4), 1013-1062.
- [27] Flinn, C and J. Mabili (2010), "On-the-Job Search, Minimum Wages, and Labor Market Outcomes in an Equilibrium Bargaining Framework", NYU mimeo.
- [28] Guiso L., L. Pistaferri and F. Schivardi (2005), "Insurance within the Firm" *Journal of Political Economy*, 113 (5), 1054-1087.
- [29] Hellerstein, J. K. and D. Neumark (1999), "Sex, Wages, and Productivity: An Empirical Analysis of Israeli Firm-Level Data". *International Economic Review* 40 (1), 95–123.
- [30] Hellerstein, J. K., D. Neumark, and K. R. Troske (1999), "Wages, Productivity, and Worker Characteristics: Evidence from Plant-Level Pro-

- duction Functions and Wage Equations". *Journal of Labor Economics* 17 (3), 409–46.
- [31] Jolivet, G., F. Postel-Vinay and J.-M. Robin (2006), "The Empirical Content of the Job Search Model: Labor Mobility and Wage Distributions in Europe and the U.S." *European Economic Review*, Vol. 50, 877-907.
 - [32] Kawaguchi D. (2007), "A market test for sex discrimination: Evidence from Japanese firm-level panel data" *International Journal of Industrial Organization* Volume 25 (3),-460
 - [33] Lauer, C. (2003), "Education and Unemployment: A French-German Comparison". ZEW Discussion Paper No. 03-34.
 - [34] Lopez-Acevedo, G., M. Tinajero and M. Rubio (2005), "Mexico: Human Capital Effects on Wages and Productivity " *World Bank Policy Research Working Paper* No. 3791
 - [35] Lindeboom, M., R. Mendes and G.J. van den Berg (2010), "An Empirical Assessment of Assortative Matching in the Labor Market" *Labour Economics*, forthcoming.
 - [36] Mondal S.(2006), "Employer Discrimination and Racial Differences in Wages and Turnover" Unpublished Paper.
 - [37] Mortensen, D.T. (2003), *Wage Dispersion*, MIT Press, Cambridge.
 - [38] Moscarini, G, (2008), "Job-to-Job Quits and Corporate Culture" Unpublished Working Paper.
 - [39] Postel-Vinay, F. and J-M Robin (2002), "Equilibrium Wage Dispersion with Heterogeneous Workers and Firms." *Econometrica* 70, 2295-1350.
 - [40] Ridder, G. and G.J. van den Berg (2003), "Measuring labor market frictions: a cross-country comparison", *Journal of the European Economic Association*, 1, 224-244.

- [41] Rubinstein, A. (1982), "Perfect Equilibrium in a Bargaining Model," *Econometrica*, 50, 97–109.
- [42] Shimer, R. (2006), "On-the-Job Search and Strategic Bargaining," *European Economic Review* 50, 811.830
- [43] Shimer, R. and L. Smith (2000), "Assortative Matching and Search" *Econometrica*, 68 (2), 343 369
- [44] Sulis G. (2007), "Gender Wage Differentials in Italy: a Structural Estimation Approach" CRENoS WP 07-15
- [45] Van Biesebroeck J. (2009) "Wage and Productivity Premiums in Sub-Saharan Africa," in S. Bender, J. Lane, K. Shaw, F. Andersson, and T. von Wachter (eds.) *The Analysis of Firms and Employees: Quantitative and Qualitative Approaches* (forthcoming) University of Chicago Press
- [46] Van den Berg, G.J. and A. van Vuuren (2003), "The Effect of Search Frictions on Wages" IZA DP No. 805
- [47] Verner, D. (1999) "Wage and Productivity Gaps: Evidence from Ghana," World Bank Policy Research Working Paper No. 2168
- [48] Wagner, G., R.V. Burkhauser, and F. Behringer (1993), "The English Language Public Use File of The German Socio-Economic Panel," *Journal of Human Resources*, 28(2), 429-433.